Exercises for Algorithms and Data Structures

Antonio Carzaniga
Faculty of Informatics
USI
(Università della Svizzera italiana)

Edition 2.8
February 2020
(with some solutions)

►**Exercise 1.** Answer the following questions on the big-oh notation.

Question 1: Explain what $g(n) = O(f(n))$ means. *(5')*

Question 2: Explain why it is meaningless to state that "the running time of algorithm A is *at least* $O(n^2)$." *(5')*

Question 3: Given two functions $f = \Omega(\log n)$ and $g = O(n)$, consider the following statements. For each statement, write whether it is true or false. For each false statement, write two functions f and g that show a counter-example. *(5')*

- $g(n) = O(f(n))$
- $f(n) = O(g(n))$
- $f(n) = \Omega(\log(g(n)))$
- $f(n) = \Theta(\log(g(n)))$
- $f(n) + g(n) = \Omega(\log n)$

Question 4: For each one of the following statements, write two functions f and g that satisfy the given condition. *(5')*

- $f(n) = O(g^2(n))$
- $f(n) = \omega(g(n))$
- $f(n) = \omega(\log(g(n)))$
- $f(n) = \Omega(f(n)g(n))$
- $f(n) = \Theta(g(n)) + \Omega(g^2(n))$

►**Exercise 2.** Write an algorithm called FIND-LARGEST that finds the largest number in an array using a divide-and-conquer strategy. Also, write the time complexity of your algorithm in terms of big-oh notation. Briefly justify your complexity analysis. *(20')*

►**Exercise 3.** Illustrate the execution of the *merge-sort* algorithm on the array

$$A = \langle 3, 13, 89, 34, 21, 44, 99, 56, 9 \rangle$$

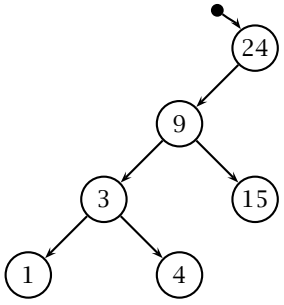
For each fundamental iteration or recursion of the algorithm, write the content of the array. Assume the algorithm performs an in-place sort. *(20')*

►**Exercise 4.** Consider the array $A = \langle 29, 18, 10, 15, 20, 9, 5, 13, 2, 4, 15 \rangle$.

Question 1: Does A satisfy the *max-heap* property? If not, fix it by swapping two elements. *(5')*

Question 2: Using array A (possibly corrected), illustrate the execution of the *heap-extract-max* algorithm, which extracts the max element and then rearranges the array to satisfy the *max-heap* property. For each iteration or recursion of the algorithm, write the content of the array A. *(15')*

►**Exercise 5.** Consider the following binary search tree (BST).



Question 1: List all the possible insertion orders (i.e., permutations) of the keys that could have produced this BST. *(5')*

Question 2: Draw the same BST after the insertion of keys: 6, 45, 32, 98, 55, and 69, in this order. *(5')*

Question 3: Draw the BST resulting from the deletion of keys 9 and 45 from the BST resulting from question 2. *(5')*

Question 4: Write at least three insertion orders (permutations) of the keys remaining in the BST after question 3 that would produce a balanced tree (i.e., a minimum-height tree). *(5')*

►**Exercise 6.** Consider a hash table that stores integer keys. The keys are 32-bit unsigned values, and are always a power of 2. Give the minimum table size t and the hash function $h(x)$ that takes a key x and produces a number between 1 and t, such that no collision occurs. *(10')*

►**Exercise 7.** Explain why the time complexity of searching for elements in a hash table, where conflicts are resolved by chaining, decreases as its load factor α decreases. Recall that α is defined as the ratio between the total number of elements stored in the hash table and the number of slots in the table.

►**Exercise 8.** For each statement below, write whether it is true or false. For each false statement, write a counter-example. *(10')*

- $f(n) = \Theta(n) \wedge g(n) = \Omega(n) \Rightarrow f(n)g(n) = \Omega(n^2)$
- $f(n) = \Theta(1) \Rightarrow n^{f(n)} = O(n)$
- $f(n) = \Omega(n) \wedge g(n) = O(n^2) \Rightarrow g(n)/f(n) = O(n)$
- $f(n) = O(n^2) \wedge g(n) = O(n) \Rightarrow f(g(n)) = O(n^3)$
- $f(n) = O(\log n) \Rightarrow 2^{f(n)} = O(n)$
- $f = \Omega(\log n) \Rightarrow 2^{f(n)} = \Omega(n)$

►**Exercise 9.** Write tight asymptotic bounds for each one of the following definitions of $f(n)$. *(10')*

- $g(n) = \Omega(n) \wedge f(n) = g(n)^2 + n^3 \Rightarrow f(n) =$
- $g(n) = O(n^2) \wedge f(n) = n \log(g(n)) \Rightarrow f(n) =$
- $g(n) = \Omega(\sqrt{n}) \wedge f(n) = g(n + 2^{16}) \Rightarrow f(n) =$
- $g(n) = \Theta(n) \wedge f(n) = 1 + 1/\sqrt{g(n)} \Rightarrow f(n) =$
- $g(n) = O(n) \wedge f(n) = 1 + 1/\sqrt{g(n)} \Rightarrow f(n) =$
- $g(n) = O(n) \wedge f(n) = g(g(n)) \Rightarrow f(n) =$

►**Exercise 10.** Write the ternary-search trie (TST) that represents a dictionary of the strings: “gnu” “emacs” “gpg” “else” “gnome” “go” “eps2eps” “expr” “exec” “google” “elif” “email” “exit” “epstopdf” *(10')*

►**Exercise 11.** Answer the following questions.

Question 1: A hash table with chaining is implemented through a table of K slots. What is the expected number of steps for a search operation over a set of $N = K/2$ keys? Briefly justify your answers.

Question 2: What are the worst-case, average-case, and best-case complexities of *insertion-sort*, *bubble-sort*, *merge-sort*, and *quicksort*? *(5')*

►**Exercise 12.** Write the pseudo code of the in-place *insertion-sort* algorithm, and illustrate its execution on the array

$$A = \langle 7, 17, 89, 74, 21, 7, 43, 9, 26, 10 \rangle$$

Do that by writing the content of the array at each main (outer) iteration of the algorithm. *(20')*

►**Exercise 13.** Consider a binary tree containing N integer keys whose values are all less than K, and the following FIND-PRIME algorithm that operates on this tree.

```
FIND-PRIME(T)
1  x = TREE-MIN(T)
2  while x ≠ NIL
3      x = TREE-SUCCESSOR(x)
4      if IS-PRIME(x.key)
5          return x
6  return x
```

```
IS-PRIME(n)
1  i = 2
2  while i · i ≤ n
3      if i divides n
4          return FALSE
5      i = i + 1
6  return TRUE
```

Hint: these are the relevant binary-tree algorithms.

```
TREE-SUCCESSOR(x)
1  if x.right ≠ NIL
2  return TREE-MINIMUM(x.right)
3  y = x.parent
4  while y ≠ NIL and x == y.right
5      x = y
6      y = y.parent
7  return y
```

```
TREE-MINIMUM(x)
1  while x.left ≠ NIL
2      x = x.left
3  return x
```

Write the time complexity of FIND-PRIME. Justify your answer. *(10')*

►**Exercise 14.** Consider the following *max-heap*

$$H = \langle 37, 12, 30, 10, 3, 9, 20, 3, 7, 1, 1, 7, 5 \rangle$$

Write the exact output of the following EXTRACT-ALL algorithm run on H

```
EXTRACT-ALL(H)
1  while H.heap-size > 0
2      HEAP-EXTRACT-MAX(H)
3      for i = 1 to H.heap-size
4          output H[i]
5      output "." END-OF-LINE
```

```
HEAP-EXTRACT-MAX(H)
1  if H.heap-size > 0
2      k = H[1]
3      H[1] = H[H.heap-size]
4      H.heap-size = H.heap-size − 1
5      MAX-HEAPIFY(H)
6  return k
```

(20')

►**Exercise 15.** Develop an efficient in-place algorithm called PARTITION-EVEN-ODD(A) that partitions an array A in *even* and *odd* numbers. The algorithm must terminate with A containing all its *even* elements preceding all its *odd* elements. For example, for input $A = \langle 7, 17, 74, 21, 7, 9, 26, 10 \rangle$, the result might be $A = \langle 74, 10, 26, 17, 7, 21, 9, 7 \rangle$. PARTITION-EVEN-ODD must be an *in-place* algorithm, which means that it may use only a constant memory space in addition to A. In practice, this means that you may not use another temporary array.

Question 1: Write the pseudo-code for PARTITION-EVEN-ODD. *(20')*

Question 2: Characterize the complexity of PARTITION-EVEN-ODD. Briefly justify your answer. *(10')*

Question 3: Formalize the correctness of the partition problem as stated above, and prove that PARTITION-EVEN-ODD is correct using a loop-invariant. *(20')*

Question 4: If the complexity of your algorithm is not already linear in the size of the array, write a new algorithm PARTITION-EVEN-ODD-OPTIMAL with complexity $O(N)$ (with $N = |A|$). *(20')*

►**Exercise 16.** The binary string below is the title of a song encoded using Huffman codes.

001100010111110110011101110110000010011101001010 1

Given the letter frequencies listed in the table below, build the Huffman codes and use them to decode the title. In cases where there are multiple "greedy" choices, the codes are assembled by combining the first letters (or groups of letters) from left to right, in the order given in the table. Also, the codes are assigned by labeling the left and right branches of the prefix/code tree with '0' and '1', respectively.

letter	a	h	v	w	‘ ’	e	t	l	o
frequency	1	1	1	1	2	2	2	3	3

(20')

►**Exercise 17.** Consider the *text* and *pattern* strings:
text: `momify my mom please`
pattern: `mom`
Use the Boyer-Moore string-matching algorithm to search for the pattern in the text. For each character comparison performed by the algorithm, write the current *shift* and highlight the character position considered in the pattern string. Assume that indexes start from 0. The following table shows the first comparison as an example. Fill the rest of the table. *(10')*

n.	*shift*	m	o	m	i	f	y		m	y		m	o	m		p	l	e	a	s	e
1	0	m	o	m																	
2																					
…	…	…																			

►**Exercise 18.** You wish to create a database of stars. For each star, the database will store several megabytes of data. Considering that your database will store billions of stars, choose the data structure that will provide the best performance. With this data structure you should be able to find, insert, and delete stars. Justify your choice. *(10')*

►**Exercise 19.** You are given a set of persons P and their friendship relation R. That is, $(a, b) \in R$ if and only if a is a friend of b. You must find a way to introduce person x to person y through a chain of friends. Model this problem with a graph and describe a strategy to solve the problem. *(10')*

►**Exercise 20.** Answer the following questions

Question 1: Explain what $f(n) = \Omega(g(n))$ means. *(5')*

Question 2: Explain what kind of problems are in the P complexity class. *(5')*

Question 3: Explain what kind of problems are in the NP complexity class. *(5')*

Question 4: Explain what it means for problem A to be *polynomially-reducible* to problem B. *(5')*

Question 5: Write *true*, *false*, or *unknown* depending on whether the assertions below are true, false, or we do not know. *(5')*

- $\mathrm{P} \subseteq \mathrm{NP}$
- $\mathrm{NP} \subseteq \mathrm{P}$
- $n! = O(n^{100})$
- $\sqrt{n} = \Omega(\log n)$
- $3n^2 + \frac{1}{n} + 4 = \Theta(n^2)$

Question 6: Consider the follwing *exact-change problem*. Given a collection of n values $V = \{v_1, v_2, \ldots, v_n\}$ representing coins and bills in a cash register, and given a value x, output 1 if there exists a subset of V whose total value is equal to x, or 0 otherwise. Is the exact-change problem in NP? Justify your answer. *(5')*

►**Exercise 21.** A thief robbing a gourmet store finds n pieces of precious cheeses. For each piece i, v_i designates its value and w_i designates its weight. Considering that W is the maximum weight the thief can carry, and considering that the thief may take any fraction of each piece, you must find the quantity of each piece the thief must take to maximize the value of the robbery. *(20')*

Question 1: Write an algorithm that solves the problem using a *greedy* or *dynamic programming* strategy. Analyze the complexity of your solution.

Question 2: Prove that the problem has an *optimal substructure*, meaning that an optimal solution to a problem instance X contains within it some optimal solutions to subproblems $Y \subseteq X$.

Question 3: Show the *greedy choice property* also holds for some greedy-choice strategy. Recall that the greedy-choice property holds if and only if every greedy choice according to the given strategy is contained in an optimal solution.

►**Exercise 22.** You are in front of a stack of pancakes of different diameter. Unfortunately, you cannot eat them unless they are sorted according to their size, with the biggest one at the bottom. To sort them, you are given a spatula that you can use to split the stack in two parts and then flip the top part of the stack. Write the an algorithm SORT-PANCAKES(P) that sorts the stack P using only spatula-flip operations. The array P stores the pancakes top-to-bottom, thus $P[1]$ is the size of the pancake at the top of the stack, while $P[P.length]$ is the size of the pancake at the bottom of the stack. Your algorithm must indicate a spatula flip with the spatula inserted at position i with SPATULA-FLIP(P, i), which flips all the elements in $P[1 \ldots i]$. *(20')*

►**Exercise 23.** The following matrix represents a directed graph over vertices $a, b, c, \ldots, \ell$. Rows and columns represent the source and destination of edges, respectively.

	a	b	c	d	e	f	g	h	i	j	k	ℓ
a					1	1						
b										1		
c								1			1	
d			1									
e		1								1		
f		1								1		
g			1	1								
h											1	1
i			1				1					
j												
k												1
ℓ												

Sort the vertices in a *reverse topological order* using the *depth-first search* algorithm. (**Hint:** if you order the vertices from left to right in reverse topological order, then all edges go from right to left.) Justify your answer by showing the relevant data maintained by the depth-first search algorithm, and by explaining how that can be used to produce a reverse topological order. *(15')*

►**Exercise 24.** Answer the following questions on the complexity classes P an NP. Justify your answers.

Question 1: P $\subseteq$ NP? *(5')*

Question 2: A problem Q is in P and there is a polynomial-time reduction from Q to Q'. What can we say about Q'? Is $Q' \in$ P? Is $Q' \in$ NP? *(5')*

Question 3: Let Q be a problem defined as follows. *Input:* a set of numbers $A = \{a_1, a_2, \ldots, a_N\}$ and a number x; *Output:* 1 if and only if there are two values $a_i, a_k \in A$ such that $a_i + a_k = x$. Is Q in NP? Is Q in P? *(5')*

►**Exercise 25.** Consider the *subset-sum* problem: given a set of numbers $A = \{a_1, a_2, \ldots, a_n\}$ and a number x, output TRUE if there is a subset of numbers in A that add up to x, otherwise output FALSE. Formally, $\exists S \subseteq A$ such that $\sum_{y \in S} y = x$. Write a dynamic-programming algorithm to solve the subset-sum problem and informally analyze its complexity. *(20')*

►**Exercise 26.** Explain the idea of *dynamic programming* using the shortest-path problem as an example. (The shortest path problem amounts to finding the shortest path in a given graph $G = (V, E)$ between two given vertices a and b.) *(15')*

►**Exercise 27.** Consider an initially empty B-Tree with minimum degree $t = 3$. Draw the B-Tree after the insertion of the keys $27, 33, 39, 1, 3, 10, 7, 200, 23, 21, 20$, and then after the additional insertion of the keys $15, 18, 19, 13, 34, 200, 100, 50, 51$. *(10')*

►**Exercise 28.** There are three containers whose sizes are 10 pints, 7 pints, and 4 pints, respectively. The 7-pint and 4-pint containers start out full of water, but the 10-pint container is initially empty. Only one type of operation is allowed: pouring the contents of one container into another, stopping only when the source container is empty, or the destination container is full. Is there a sequence of pourings that leaves exactly two pints in either the 7-pint or the 4-pint container?

Question 1: Model this as a graph problem: give a precise definition of the graph involved (type of the graph, labels on vertices, meaning of an edge). Provide the set of all reachable vertices, identify the initial vertex and the goal vertices. (**Hint:** all vertices that satisfy the condition imposed by the problem are reachable, so you don't have to draw a graph.)

Question 2: State the specific question about this graph that needs to be answered?

Question 3: What algorithm should be applied to solve the problem? Justify your answer. *(15')*

►**Exercise 29.** Write an algorithm called MOVETOROOT(x, k) that, given a binary tree rooted at node x and a key k, moves the node containing k to the root position and returns that node if k is in the tree. If k is not in the tree, the algorithm must return x (the original root) without modifying the tree. Use the typical notation whereby $x.key$ is the key stored at node x, $x.left$ and $x.right$ are the left and right children of x, respectively, and $x.parent$ is x's parent node. *(15')*

►**Exercise 30.** Given a sequence of numbers $A = \langle a_1, a_2, \ldots, a_n \rangle$, an *increasing subsequence* is a sequence $a_{i_1}, a_{i_2}, \ldots, a_{i_k}$ of elements of A such that $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, and such that $a_{i_1} < a_{i_2} < \ldots < a_{i_k}$. You must find the *longest increasing subsequence*. Solve the problem using dynamic programming.

Question 1: Define the *subproblem structure* and the solution of each subproblem. *(5')*

Question 2: Write an iterative algorithm that solves the problem. Illustrate the execution of the algorithm on the sequence $A = \langle 2, 4, 5, 6, 7, 9 \rangle$. *(10')*

Question 3: Write a recursive algorithm that solves the problem. Draw a tree of recursive calls for the algorithm execution on the sequence $A = \langle 1, 2, 3, 4, 5 \rangle$. *(10')*

Question 4: Compare the time complexities of the iterative and recursive algorithms. *(5')*

►**Exercise 31.** One way to implement a *disjoint-set* data structure is to represent each set by a linked list. The first node in each linked list serves as the representative of its set. Each node contains a key, a pointer to the next node, and a pointer back to the representative node. Each list maintains the pointers *head*, to the representative, and *tail*, to the last node in the list.

Question 1: Write the pseudo-code and analyze the time complexity for the following operations:

- MAKE-SET(x): creates a new set whose only member is x.
- UNION(x, y): returns the representative of the union of the sets that contain x and y.
- FIND-SET(x): returns a pointer to the representative of the set containing x.

Note that x and y are nodes. *(15')*

Question 2: Illustrate the linked list representation of the following sets:

- $\{c, a, d, b\}$
- $\{e, g, f\}$
- UNION(d, g)

(5')

►**Exercise 32.** Explain what it means for a hash function to be perfect for a given set of keys. Consider the hash function $h(x) = x \mod m$ that maps an integer x to a table entry in $\{0, 1, \ldots m-1\}$. Find an $m \leq 12$ such that h is a perfect hash function on the set of keys $\{0, 6, 9, 12, 22, 31\}$. *(10')*

►**Exercise 33.** Draw the binary search tree obtained when the keys $1, 2, 3, 4, 5, 6, 7$ are inserted in the given order into an initially empty tree. What is the problem of the tree you get? Why is it a problem? How could you modify the insertion algorithm to solve this problem. Justify your answer. *(10')*

►**Exercise 34.** Consider the following array:

$$A = \langle 4, 33, 6, 90, 33, 32, 31, 91, 90, 89, 50, 33 \rangle$$

Question 1: Is A a *min-heap*? Justify your answer by briefly explaining the *min-heap* property. *(10')*

Question 2: If A is a *min-heap*, then extract the minimum value and then rearrange the array with the *min-heapify* procedure. In doing that, show the array at every iteration of *min-heapify*. If A is not a *min-heap*, then rearrange it to satisfy the *min-heap* property. *(10')*

►**Exercise 35.** Write the pseudo-code of the *insertion-sort* algorithm. Illustrate the execution of the algorithm on the array $A = \langle 3, 13, 89, 34, 21, 44, 99, 56, 9 \rangle$, writing the intermediate values of A at each iteration of the algorithm. *(20')*

►**Exercise 36.** Encode the following sentence with a Huffman code

Common sense is the collection of prejudices acquired by age eighteen

Write the complete construction of the code. *(20')*

►**Exercise 37.** Consider the *text* and *query* strings:
text: `It ain't over till it's over.`
query: `over`
Use the Boyer-Moore string-matching algorithm to search for the query in the text. For each character comparison performed by the algorithm, write the current *shift* and highlight the character position considered in the query string. Assume that indexes start from 0. The following table shows the first comparison as an example. Fill the rest of the table. *(10')*

n.	*shift*	I	t		a	i	n	'	t		o	v	e	r		t	i	l	l		i	t	'	s		o	v	e	r	.
1	0	o	v	e	r																									
2																														
...	...	...																												

►**Exercise 38.** Briefly answer the following questions

Question 1: What does $f(n) = \Theta(g(n))$ mean? *(5')*

Question 2: What kind of problems are in the P class? Give an example of a problem in P. *(5')*

Question 3: What kind of problems are in the NP class? Give an example of a problem in NP. *(5')*

Question 4: What does it mean for a problem A to be *reducible* to a problem B? *(5')*

►**Exercise 39.** For each of the following assertions, write "true," "false," or "?" depending on whether the assertion is true, false, or it may be either true or false. *(10')*

Question 1: $\mathrm{P} \subseteq \mathrm{NP}$

Question 2: The *knapsack* problem is in P

Question 3: The *minimal spanning tree* problem is in NP

Question 4: $n! = O(n^{100})$

Question 5: $\sqrt{n} = \Omega(\log(n))$

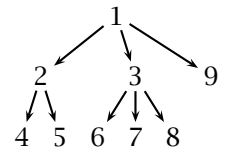
Question 6: insertion-sort performs like *quicksort* on an almost sorted sequence

►**Exercise 40.** An application must read a long sequence of numbers given in no particular order, and perform many searches on that sequence. How would you implement that application to minimize the overall time-complexity? Write exactly what algorithms you would use, and in what sequence. In particular, write the high-level structure of a *read* function, to read and store the sequence, and a *find* function too look up a number in the sequence. *(10')*

►**Exercise 41.** Write an algorithm that takes a set of (x, y) coordinates representing points on a plane, and outputs the coordinates of two points with the maximal distance. The signature of the algorithm is MAXIMAL-DISTANCE(X, Y), where X and Y are two arrays of the same length representing the x and y coordinates of each point, respectively. Also, write the asymptotic complexity of MAXIMAL-DISTANCE. Briefly justify your answer. *(10')*

►**Exercise 42.** A *directed tree* is represented as follows: for each vertex v, $v.$*first-child* is either the first element in a list of child-vertices, or NIL if v is a leaf. For each vertex v, $v.$*next-sibling* is the next element in the list of v's siblings, or NIL if v is the last element in the list. For example, the arrays on the left represent the tree on the right:

v	1	2	3	4	5	6	7	8	9
first-child	2	4	6	NIL	NIL	NIL	NIL	NIL	NIL
next-sibling	NIL	3	9	5	NIL	7	8	NIL	NIL



Question 1: Write two algorithms, MAX-DEPTH(*root*) and MIN-DEPTH(*root*), that, given a tree, return the maximal and minimal depth of any leaf vertex, respectively. (E.g., the results for the example tree above are 2 and 1, respectively.) *(15')*

Question 2: Write an algorithm DEPTH-FIRST-ORDER(*root*) that, given a tree, prints the vertices in depth-first visitation order, such that a vertices is always preceded by all its children (e.g., the result for the example tree above is $4, 5, 2, 6, 7, 8, 3, 9, 1$). *(10')*

Question 3: Analyze the complexity of MAX-DEPTH, MIN-DEPTH and DEPTH-FIRST-ORDER. *(5')*

►**Exercise 43.** Write an algorithm called IN-PLACE-SORT(A) that takes an array of numbers, and sorts the array *in-place*. That is, using only a constant amount of extra memory. Also, give an informal analysis of the asymptotic complexity of your algorithm. *(10')*

►**Exercise 44.** Given a sequence $A = \langle a_1, \ldots, a_n \rangle$ of numbers, the *zero-sum-subsequence* problem amounts to deciding whether A contains a subsequence of consecutive elements $a_i, a_{i+1}, \ldots, a_k$, with $1 \le i \le k \le n$, such that $a_i + a_{i+1} + \cdots + a_k = 0$. Model this as a dynamic-programming problem and write a dynamic-programming algorithm ZERO-SUM-SEQUENCE(A) that, given an array A, returns TRUE if A contains a zero-sum subsequence, or FALSE otherwise. Also, give an informal analysis of the complexity of ZERO-SUM-SEQUENCE. *(30')*

►**Exercise 45.** Give an example of a randomized algorithm derived from a deterministic algorithm. Explain why there is an advantage in using the randomized variant. *(10')*

►**Exercise 46.** Implement a TERNARY-TREE-SEARCH(x, k) algorithm that takes the root of a ternary tree and returns the node containing key k. A ternary tree is conceptually identical to a binary tree, except that each node x has two keys, $x.key_1$ and $x.key_2$, and three links to child nodes, $x.left$, $x.center$, and $x.right$, such that the left, center, and right subtrees contains keys that are, respectively, less than $x.key_1$, between $x.key_1$ and $x.key_2$, and greater than $x.key_2$. Assume there are no duplicate keys. Also, assuming the tree is balanced, what is the asymptotic complexity of the algorithm? *(10')*

►**Exercise 47.** Answer the following questions. Briefly justify your answers.

Question 1: A hash table that uses chaining has M slots and holds N keys. What is the expected complexity of a search operation? *(5')*

Question 2: The asymptotic complexity of algorithm A is $\Omega(N \log N)$, while that of B is $\Theta(N^2)$. Can we compare the two algorithms? If so, which one is asymptotically faster? *(5')*

Question 3: What is the difference between "Las Vegas" and "Monte Carlo" randomized algorithms? *(5')*

Question 4: What is the main difference between the Knuth-Morris-Pratt algorithm and Boyer-Moore string-matching algorithms in terms of complexity? Which one as the best worst-case complexity? *(5')*

►**Exercise 48.** Consider *quick-sort* as an in-place sorting algorithm.

Question 1: Write the pseudo-code using only *swap* operations to modify the input array. *(10')*

Question 2: Apply the algorithm of question 1 to the array $A = \langle 8, 2, 12, 17, 4, 8, 7, 1, 12 \rangle$. Write the content of the array after each swap operation. *(10')*

►**Exercise 49.** Consider this *minimal vertex cover* problem: given a graph $G = (V, E)$, find a minimal set of vertices S such that for every edge $(u, v) \in E$, u or v (or both) are in S.

Question 1: Model *minimal vertex cover* as a dynamic-programming problem. Write the pseudo-code of a dynamic-programming solution. *(15')*

Question 2: Do you think that your model of *minimal vertex cover* admits a greedy choice? Try at least one meaningful greedy strategy. Show that it does not work, giving a counter-example graph for which the strategy produces the wrong result. (**Hint:** one meaningful strategy is to choose a maximum-degree vertex first. The degree of a vertex is the number of its incident edges.) *(5')*

►**Exercise 50.** The graph $G = (V, E)$ represents a social network in which each vertex represents a person, and an edge $(u, v) \in E$ represents the fact that u and v know each other. Your problem is to organize the largest party in which nobody knows each other. This is also called the *maximal independent set* problem. Formally, given a graph $G = (V, E)$, find a set of vertices S of maximal size in which no two vertices are adjacent. (I.e., for all $u \in S$ and $v \in S$, $(u, v) \notin E$.)

Question 1: Formulate a decision variant of *maximal independent set*. Say whether the problem is in NP, and briefly explain what that means. *(10')*

Question 2: Write a verification algorithm for the *maximal independent set* problem. This algorithm, called TESTINDEPENDENTSET(G, S), takes a graph G represented through its adjacency matrix, and a set S of vertices, and returns TRUE if S is a valid independent set for G. *(10')*

►**Exercise 51.** A *Hamilton cycle* is a cycle in a graph that touches every vertex exactly once. Formally, in $G = (V, E)$, an ordering of *all* vertices $H = v_1, v_2, \ldots, v_n$ forms a Hamilton cycle if $(v_n, v_1) \in E$, and $(v_i, v_{i+1}) \in E$ for all i between 1 and $n - 1$. Deciding whether a given graph is *Hamiltonian* (has a Hamilton cycle) is a well known NP-complete problem.

Question 1: Write a verification algorithm for the *Hamiltonian graph* problem. This algorithm, called TESTHAMILTONCYCLE(G, H), takes a graph G represented through adjacency lists, and an array of vertices H, and returns TRUE if H is a valid Hamilton cycle in G. *(10')*

Question 2: Give the asymptotic complexity of your implementation of TESTHAMILTONCYCLE. *(5')*

Question 3: Explain what it means for a problem to be NP-complete. *(5')*

►**Exercise 52.** Consider using a b-tree with minimum degree $t = 2$ as an in-memory data structure to implement dynamic sets.

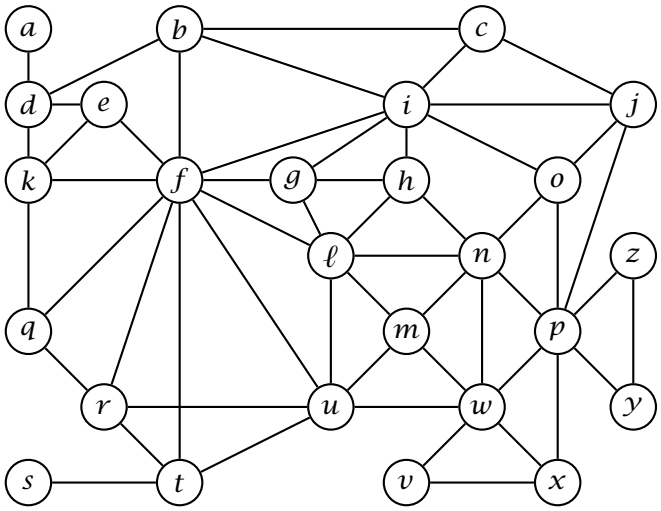
Question 1: Compare this data structure with a red-black tree. Is this data structure better, worse, or the same as a red-black tree in terms of time complexity? Briefly justify your answer. In particular, characterize the complexity of insertion and search. *(10')*

Question 2: Write an iterative (i.e., non-recursive) *search* algorithm for this degree-2 b-tree. Remember that the data structure is *in-memory*, so there is no need to perform any disk read/write operation. *(10')*

Question 3: Write the data structure after the insertion of keys $10, 3, 8, 21, 15, 4, 6, 19, 28, 31$, in this order, and then after the insertion of keys $25, 33, 7, 1, 23, 35, 24, 11, 2, 5$. *(10')*

Question 4: Write the insertion algorithm for this degree-2 b-tree. (**Hint:** since the minimum degree is fixed at 2, the insertion algorithm may be implemented in a simpler fashion without all the loops of the full b-tree insertion.) *(15')*

►**Exercise 53.** Consider a breadth-first search (BFS) on the following graph, starting from vertex a.



Write the two vectors π (previous) and d (distance), resulting from the BFS algorithm. *(10')*

►**Exercise 54.** Write a sorting algorithm that runs with in time $O(n \log n)$ in the average case (on an input array of size n). Also, characterize the best- and worst-case complexity of your solution. *(20')*

►**Exercise 55.** The following algorithms take an array A of integers. For each algorithm, write the asymptotic, best- and worst-case complexities as functions of the size of the input $n = |A|$. Your characterizations should be as tight as possible. Justify your answers by writing a short explanation of what each algorithm does. *(20')*

ALGORITHM-I(A)

```
1  for i = |A| downto 2
2      s = TRUE
3      for j = 2 to i
4          if A[j − 1] > A[j]
5              swap A[j − 1] ↔ A[j]
6              s = FALSE
7      if s == TRUE
8          return
```

ALGORITHM-II(A)

```
1  i = 1
2  j = |A|
3  while i < j
4      if A[i] > A[j]
5          swap A[i] ↔ A[i + 1]
6          if i + 1 < j
7              swap A[i] ↔ A[j]
8          i = i + 1
9      else j = j − 1
```

►**Exercise 56.** The following algorithms take a binary search tree T containing n keys. For each algorithm, write the asymptotic, best- and worst-case complexities as functions of n. Your characterizations should be as tight as possible. Justify your answers by writing a short explanation of what each algorithm does. *(20')*

```
ALGORITHM-III(T, k)
 1  if T == NIL
 2       return FALSE
 3  if T.key == k
 4       return TRUE
 5  if ALGORITHM-III(T.left)
 6       return TRUE
 7  else return ALGORITHM-III(T.right)
```

```
ALGORITHM-IV(T, k_1, k_2)
 1  if T == NIL
 2       return 0
 3  if k_1 > k_2
 4       swap k_1 ↔ k_2
 5  r = 0
 6  if T.key < k_2
 7       r = r + ALGORITHM-IV(T.right, k_1, k_2)
 8  if T.key > k_1
 9       r = r + ALGORITHM-IV(T.left, k_1, k_2)
10  if T.key < k_2 and T.key > k_1
11       r = r + 1
12  return r
```

►**Exercise 57.** Answer the following questions on complexity theory. Justify your answers. All problems are decision problems. (*Hint:* answers are not limited to "yes" or "no.") *(20')*

Question 1: An algorithm A solves a problem P of size n in time $O(n^3)$. Is P in NP?

Question 2: An algorithm A solves a problem P of size n in time $\Omega(n \log n)$. Is P in P? Is it in NP?

Question 3: A problem P in NP can be polynomially reduced into a problem Q. Is Q in P? Is Q in NP?

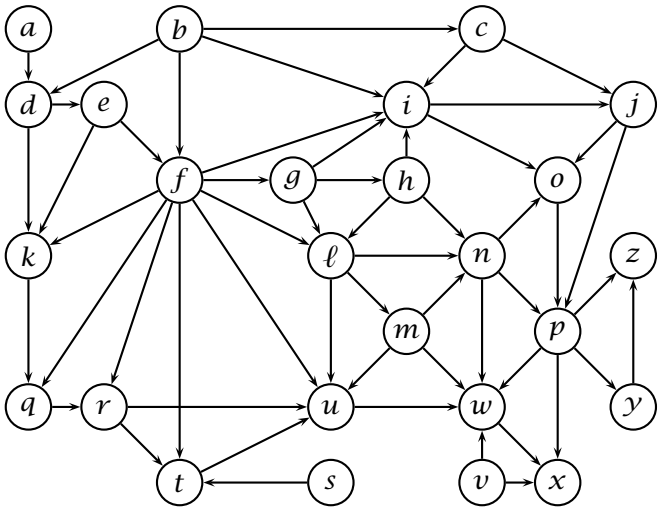
Question 4: A problem P can be polynomially reduced into a problem Q in NP. Is P in P? Is P NP-hard?

Question 5: A problem P of size n does not admit to any algorithmic solution with complexity $O(2^n)$. Is P in P? Is P in NP?

Question 6: An algorithm A takes an instance of a problem P of size n and a "certificate" of size $O(n^c)$, for some constant c, and *verifies* in time $O(n^2)$ that the solution to given problem is affirmative. Is P in P? Is P in NP? Is P NP-complete?

►**Exercise 58.** Write an algorithm TSTCOUNTGREATER(T, s) that takes the root T of a ternary-search trie (TST) and a string s, and returns the number of strings stored in the trie that are lexicographically greater than s. Given a node T, $T.left$, $T.middle$, and $T.right$ are the left, middle, and right subtrees, respectively; $T.value$ is the value stored in T. The TST uses the special character '#' as the string terminator. Given two characters a and b, the relation $a < b$ defines the lexicographical order, and the terminator character is *less than* every other character. (**Hint:** first write an algorithm that, given a tree (node) counts *all* the strings stored in that tree.) *(20')*

►**Exercise 59.** Consider a depth-first search (DFS) on the following graph.



Write the three vectors π, d, and f that, for each vertex represent the *previous* vertex in the depth-first forest, the *discovery* time, and the *finish* time, respectively. Whenever necessary, iterate through vertexes in alphabetic order. *(20')*

►**Exercise 60.** Write an implementation of a *radix tree* in Java. The tree must store 32-bit integer keys. Each node in the tree must contain at most 256 links or keys, so each node would cover at most 8 bits of the key. You must implement a class called `RadixTree` with two methods, `void insert(int k)` and `boolean find(int k)`. You must also specify every other class you might use. For example, you would probably want to define a class `RadixTreeNode` to represent nodes in the tree. *(20')*

►**Exercise 61.** Answer the following questions about *red-black trees*.

Question 1: Describe the structure and the properties of a red-black tree. *(5')*

Question 2: Write an algorithm RB-TREE-SEARCH(T, k) that, given a red-black tree T and a key k, returns TRUE if T contains key k, or FALSE otherwise. *(5')*

Question 3: Let h be the height of a red-black tree containing n keys, prove that

$$h \leq 2\log(n+1)$$

Hint: first give an outline of the proof. Even if you can not give a complete proof, try to explain informally how the red-black tree property limits the height of a red-black tree. *(10')*

►**Exercise 62.** Sort the following functions in ascending order of asymptotic growth rate:

$$f_1(n) = 3^n \qquad f_2(n) = n^{1/3} \qquad f_3(n) = \log^2 n$$
$$f_4(n) = n^{\log n} \qquad f_5(n) = n^3 \qquad f_6(n) = 4^{\log n}$$
$$f_7(n) = n^2\sqrt{n} \qquad f_8(n) = 2^{2n} \qquad f_9(n) = \sqrt{\log n}$$

That is, write the sequence of sorted indexes $a_1, a_2, \ldots, a_9$ such that for all indexes a_i, a_j with $i < j$, $f_{a_i}(n) = O(f_{a_j}(n))$. (Notice that $\log n$ means $\log_2 n$.) *(10')*

►**Exercise 63.** Consider the following algorithm:

```
ALGO-A(X)
1  d = ∞
2  for i = 1 to X.length − 1
3      for j = i + 1 to X.length
4          if |X[i] − X[j]| < d
5              d = |X[i] − X[j]|
6  return d
```

Question 1: Interpreting X as an array of coordinates of points on the x-axis, explain concisely what algorithm ALGO-A does, and give a tight asymptotic bound for the complexity of ALGO-A. *(5')*

Question 2: Write an algorithm BETTER-A(X) that is functionally equivalent to ALGO-A(X), but with a better asymptotic complexity. *(15')*

►**Exercise 64.** The following defines a *ternary search trie* (TST) for character strings, in Java and in pseudo-code notation:

```
class TSTNode {
    char c;                 x.c          character at node x
    boolean have_key;       x.have-key   TRUE if node x represents a key
    TSTNode left;           x.left       left child of node x
    TSTNode middle;         x.middle     middle child of node x
    TSTNode right;          x.right      right child of node x
}
```

Write an algorithm, `void TSTPrint(TSTNode t)` in Java or TST-PRINT(x) in pseudo-code that, given the root of a TST, prints all its keys in alphabetical order. *(20')*

►**Exercise 65.** A set of keys is stored in a *max-heap* H and in a *binary search tree* T. Which data structure offers the most efficient algorithm to output all the keys in descending order? Or are the two equivalent? Write both algorithms. Your algorithms may change the data structures. *(20')*

►**Exercise 66.** Answer the following questions. Briefly justify your answers. *(10')*

Question 1: Let A be an array of numbers sorted in descending order. Does A represent a max-heap (with $A.\textit{heap-size} = A.\textit{length}$)?

Question 2: A hash table has T slots and uses chaining to resolve collisions. What are the worst-case and average-case complexities of a search operation when the hash table contains N keys?

Question 3: A hash table with 9 slots, uses chaining to resolve collision, and uses the hash function $h(k) = k \mod 9$ (slots are numbered $0, \ldots, 8$). Draw the hash table after the insertion of keys 5, 28, 19, 15, 20, 33, 12, 17, and 10.

Question 4: Is the operation of deletion in a binary search tree *commutative* in the sense that deleting x and then y from a binary search tree leaves the same tree as deleting y and then x? Argue why it is or give a counter-example.

►**Exercise 67.** Draw a binary search tree containing keys $8, 27, 13, 15, 32, 20, 12, 50, 29, 11$, inserted in this order. Then, add keys $14, 18, 30, 31$, in this order, and again draw the tree. Then delete keys 29 and 27, in this order, and again draw the tree. *(10')*

►**Exercise 68.** Consider a *max-heap* containing keys $8, 27, 13, 15, 32, 20, 12, 50, 29, 11$, inserted in this order in an initially empty heap. Write the content of the array that stores the heap. Then, insert keys 43 and 51, and again write the content of the array. Then, extract the maximum value three times, and again write the content of the array. In all three cases, write the heap as an array. *(10')*

►**Exercise 69.** Consider a *min-heap* H and the following algorithm.

BST-FROM-MIN-HEAP(H)
1 T = NEW-EMPTY-TREE()
2 **for** $i = 1$ **to** $H.\textit{heap-length}$
3 TREE-INSERT($T, H[i]$) // binary-search-tree insertion
4 **return** T

Prove that BST-FROM-MIN-HEAP does not always produce minimum-height binary trees. *(10')*

►**Exercise 70.** Consider an array A containing n numbers and satisfying the *min-heap* property. Write an algorithm MIN-HEAP-FAST-SEARCH(A, k) that finds k in A with a time complexity that is better than linear in n whenever at most $\sqrt{n}$ of the values in A are less than k. *(20')*

►**Exercise 71.** Write an algorithm B-TREE-TOP-K(R, k) that, given the root R of a b-tree of minimum degree t, and an integer k, outputs the largest k keys in the b-tree. You may assume that the entire b-tree resides in main memory, so no disk access is required. Recall that a node x in a b-tree has

the following properties: $x.n$ is the number of keys, $X.key[1] \leq x.key[2] \leq \ldots x.key[x.n]$ are the keys, $x.leaf$ tells whether x is a leaf, and $x.c[1], x.c[2], \ldots, x.c[x.n+1]$ are the pointers to x's children. *(30')*

►**Exercise 72.** Your computer has a special machine instruction called SORT-FIVE(A, i) that, given an array A and a position i, sorts in-place and in a single step the elements $A[i \ldots i+5]$ (or $A[i \ldots |A|]$ if $|A| < i+5$). Write an in-place sorting algorithm called SORT-WITH-SORT-FIVE that uses only SORT-FIVE to modify the array A. Also, analyze the complexity of SORT-WITH-SORT-FIVE. *(20')*

►**Exercise 73.** For each of the following statements, briefly argue why they are true, or show a counter-example. *(10')*

Question 1: $f(n) = O(n!) \Longrightarrow \log(f(n)) = O(n \log n)$

Question 2: $f(n) = \Theta(f(n/2))$

Question 3: $f(n) + g(n) = \Theta(\min(f(n), g(n)))$

Question 4: $f(n)g(n) = O(\max(f(n), g(n)))$

Question 5: $f(g(n)) = \Omega(\min(f(n), g(n)))$

►**Exercise 74.** Analyze the complexity of the following algorithm. *(10')*

```
SHUFFLE-A-BIT(A)
 1  i = 1
 2  j = A.length
 3  if j > i
 4      while j > i
 5          p = CHOOSE-UNIFORMLY({0, 1})
 6          if p == 1
 7              swap A[i] ↔ A[j]
 8          j = j − 1
 9          i = i + 1
10      SHUFFLE-A-BIT(A[1 ... j])
11      SHUFFLE-A-BIT(A[i ... A.length])
```

►**Exercise 75.** Answer the following questions. For each question, write "yes" when the answer is always true, "no" when it is always false, "undefined" when it can be true or false. *(10')*

Question 1: Algorithm A solves decision problem X in time $O(n \log n)$. Is X in NP?

Question 2: Is X in P?

Question 3: Decision problem X in P can be polynomially reduced to problem Y. Is there a polynomial-time algorithm to solve Y?

Question 4: Decision problem X can be polynomially reduced to a problem Y for which there is a polynomial-time verification algorithm. Is X in NP?

Question 5: Is X in P?

Question 6: An NP-hard decision problem X can be polynomially reduced to problem Y. Is Y in NP?

Question 7: Is Y NP-hard?

Question 8: Algorithm A solves decision problem X in time $\Theta(2^n)$. Is X in NP?

Question 9: Is X in P?

►**Exercise 76.** Write a minimal character-based binary code for the following sentence:

in theory, there is no difference between theory and practice; in practice, there is.

The code must map each character, including spaces and punctuation marks, to a binary string so that the total length of the encoded sentence is minimal. Use a Huffman code and show the derivation of the code. *(20')*

►**Exercise 77.** The following matrix represents a directed graph over 12 vertices labeled $a, b, \ldots, \ell$. Rows and columns represent the source and destination of edges, respectively. For example, the value 1 in row a and column f indicates an edge from a to f.

	a	b	c	d	e	f	g	h	i	j	k	ℓ
a						1						
b								1	1	1		
c									1		1	
d	1		1		1							1
e							1			1	1	
f					1					1	1	1
g		1										
h		1		1					1	1		1
i								1				
j		1					1	1				
k	1							1		1		
ℓ									1		1	

Run a *breadth-first search* on the graph starting from vertex a. Using the table below, write the two vectors π (previous) and d (distance) at each main iteration of the BFS algorithm. Write the pair π, d in each cell; for each iteration, write only the values that change. Also, write the complete BFS tree after the termination of the algorithm. *(20')*

a	b	c	d	e	f	g	h	i	j	k	ℓ
$a, 0$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$	$-, \infty$

►**Exercise 78.** A *graph coloring* associates a color with each vertex of a graph so that adjacent vertices have different colors. Write a greedy algorithm that tries to color a given graph with the least number of colors. This is a well known and difficult problem for which, most likely, there is no perfect greedy strategy. So, you should use a *reasonable* strategy, and it is okay if your algorithm does not return the absolute best coloring. The result must be a *color* array, where $v.color$ is a number representing the color of vertex v. Write the algorithm, analyze its complexity, and also show an example in which the algorithm does not achieve the best possible result. *(20')*

►**Exercise 79.** Given an array A and a positive integer k, the *selection* problem amounts to finding the largest element $x \in A$ such that at most k elements of A are less than or equal to x, or NIL if no such element exists. A simple way to implement it is as follows:

SIMPLESELECTION(A, k)

```
1  if k > A.length
2        return NIL
3  else sort A in ascending order
4        return A[k]
```

Write another algorithm that solves the selection problem without first sorting A. (**Hint:** use a divide-and-conquer strategy that "divides" A using one of its elements.) Also, illustrate the execution of the algorithm on the following input by writing its state at each main iteration or recursion.

$$A = \langle 29, 28, 35, 20, 9, 33, 8, 9, 11, 6, 21, 28, 18, 36, 1 \rangle \qquad k = 6$$

(20')

►**Exercise 80.** Consider the following *maximum-value contiguous subsequence* problem: given a sequence of numbers $A = \langle a_1, a_2, \ldots, a_n \rangle$, find two positions i and j, with $1 \le i \le j \le n$, such that the sum $a_i + a_{i+1} + \cdots + a_j$ is maximal.

Question 1: Write an algorithm to solve the problem and analyze its complexity. *(10')*

Question 2: If you have not already done so for question 1, write an algorithm that solves the maximum-value contiguous subsequence problem in time $O(n)$. (**Hint:** one such algorithm uses dynamic-programming.) *(20')*

►**Exercise 81.** Consider the following intuitive definition of the *size* of a binary search (sub)tree t: $size(t) = 0$ if t is NIL, or $size(t) = 1 + size(t.left) + size(t.right)$ otherwise. For each node t in a tree, let attribute $t.size$ denote the size of the subtree rooted at t.

Question 1: Prove that, if for each node t in a tree T, $\max\{size(t.left), size(t.right)\} \leq \frac{2}{3} size(t)$, then the height of T is $O(\log n)$, where $n = size(T)$. *(10')*

Question 2: Write the rotation procedures ROTATE-LEFT(t) and ROTATE-RIGHT(t) that return the left- and right rotation of tree t maintaining the correct *size* attributes. *(10')*

Question 3: Write an algorithm called SELECTION(T, i) that, given a tree T where each node t carries its size in $t.size$, returns the i-th key in T. *(10')*

Question 4: A tree T is *perfectly balanced* when $\max\{size(t.left), size(t.right)\} = \lfloor size(t)/2 \rfloor$ for all nodes $t \in T$. Write an algorithm called BALANCE(T) that, using the rotation procedures defined in question 2, balances T perfectly. (**Hint:** the essential operation is to move the median value of a subtree to the root of that subtree.) *(30')*

►**Exercise 82.** Write the *heap-sort* algorithm and illustrate its execution on the following sequence.

$$A = \langle 1, 1, 24, 8, 3, 36, 34, 23, 4, 30 \rangle$$

Assuming the sequence A is stored in an array passed to the algorithm, for each main iteration (or recursion) of the algorithm, write the content of the array. *(10')*

►**Exercise 83.** A radix tree is used to represent a dictionary of words defined over the alphabet of the 26 letters of the English language. Assume that letters from A to Z are represented as numbers from 1 to 26. For each node x of the tree, $x.links$ is the array of links to other nodes, and $x.value$ is a Boolean value that is true when x represents a word in the dictionary. Write an algorithm PRINT-RADIX-TREE(T) that outputs all the words in the dictionary rooted at T. *(10')*

►**Exercise 84.** Consider the following algorithm that takes an array A of length $A.length$:

```
ALGO-X(A)
1  for i = 3 to A.length
2      for j = 2 to i − 1
3          for k = 1 to j − 1
4              if |A[i] − A[j]| == |A[j] − A[k]|
               or |A[i] − A[k]| == |A[k] − A[j]|
               or |A[k] − A[i]| == |A[i] − A[j]|
5                  return TRUE
6  return FALSE
```

Write an algorithm BETTER-ALGO-X(A) equivalent to ALGO-X(A) (for all A) but with a strictly better asymptotic complexity than ALGO-X(A). *(20')*

►**Exercise 85.** For each of the following statements, write whether it is correct or not. Justify your answer by briefly arguing why it is correct, or otherwise by giving a counter example. *(10')*

Question 1: If $f(n) = O(g^2(n))$ then $f(n) = \Omega(g(n))$.

Question 2: If $f(n) = \Theta(2^n)$ then $f(n) = \Theta(3^n)$.

Question 3: If $f(n) = O(n^3)$ then $\log(f(n)) = O(\log n)$.

Question 4: $f(n) = \Theta(f(2n))$

Question 5: $f(2n) = \Omega(f(n))$

►**Exercise 86.** Write an algorithm PARTITION(A, k) that, given an array A of numbers and a value k, changes A in-place by only swapping two of its elements at a time so that all elements that are less then or equal to k precede all other elements. *(10')*

►**Exercise 87.** Consider an initially empty B-Tree with minimum degree $t = 2$.

Question 1: Draw the tree after the insertion of keys 81, 56, 16, 31, 50, 71, 58, 83, 0, and 60 in this order. *(10')*

Question 2: Can a different insertion order produce a different tree? If so, write the same set of keys in a different order and the corresponding B-Tree. If not, explain why. *(10')*

►**Exercise 88.** Consider the following decision problem. Given a set of integers A, output 1 if some of the numbers in A add up to a multiple of 10, or 0 otherwise.

Question 1: Is this problem in NP? If it is, then write a corresponding verification algorithm. If not, explain why not. *(5')*

Question 2: Is this problem in P? If it is, then write a polynomial-time solution algorithm. Otherwise, argue why not. (*Hint:* consider the input values modulo 10. That is, for each input value, consider the remainder of its division by 10.) *(15')*

►**Exercise 89.** The following greedy algorithm is intended to find the shortest path between vertices u and v in a graph $G = (V, E, w)$, where $w(x, y)$ is the length of edge $(x, y) \in E$.

GREEDY-SHORTEST-PATH$(G = (V, E, w), u, v)$

```
1  Visited = {u}                    // this is a set
2  path = ⟨u⟩                       // this is a sequence
3  while path not empty
4      x = last vertex in path
5      if x == v
6          return path
7      y = vertex y ∈ Adj[x] such that y ∉ Visited and w(x, y) is minimal
                                    // y is x's closest neighbor not already visited
8      if y == UNDEFINED            // all neighbors of x have already been visited
9          path = path − ⟨x⟩        // removes the last element y from path
10     else Visited = Visited ∪ {y}
11         path = path + ⟨y⟩        // append y to path
12 return UNDEFINED                 // there is no path between u and v
```

Does this algorithm find the shortest path always, sometimes, or never? If it always works, then explain its correctness by defining a suitable invariant for the main loop, or explain why the greedy choice is correct. If it works sometimes (but not always) show a positive example and a negative example, and briefly explain why the greedy choice does not work. If it is never correct, show an example and briefly explain why the greedy choice does not work. *(20')*

►**Exercise 90.** Write the quick-sort algorithm as a deterministic in-place algorithm, and then apply it to the array

$$\langle 50, 47, 92, 78, 76, 7, 60, 36, 59, 30, 50, 43 \rangle$$

Show the application of the algorithm by writing the content of the array after each main iteration or recursion. *(20')*

►**Exercise 91.** Consider an undirected graph G of n vertices represented by its adjacency matrix A. Write an algorithm called IS-CYCLIC(A) that, given the adjacency matrix A, returns TRUE if G contains a cycle, or FALSE if G is acyclic. Also, give a precise analysis of the complexity of your algorithm. *(20')*

►**Exercise 92.** A palindrome is a sequence of characters that is identical when read left-to-right and right-to-left. For example, the word "racecar" is a palindrome, as is the phrase "rats live on no evil star." Write an algorithm called LONGEST-PALINDROME(T) that, given an array of characters T, prints the longest palindrome in T, or any one of them if there are more than one. For example, if T is the text "radar radiations" then your algorithm should output "dar rad". Also, give a precise analysis of the complexity of your algorithm. *(20')*

►**Exercise 93.** Write an algorithm called OCCURRENCES that, given an array of numbers A, prints all the distinct values in A each followed by its number of occurrences. For example, if $A = \langle 28, 1, 0, 1, 0, 3, 4, 0, 0, 3 \rangle$, the algorithm should output the following five lines (here separated by a semicolon) "28 1; 1 2; 0 4; 3 2; 4 1". The algorithm may modify the content of A, but may not use any other memory. Each distinct value must be printed exactly once. Values may be printed in any order. The complexity of the algorithm must be $o(n^2)$, that is, strictly lower than $O(n^2)$. *(20')*

►**Exercise 94.** The following algorithm takes an array of line segments. Each line segment s is defined by its two end-points $s.a$ and $s.b$, each defined by their Cartesian coordinates $(s.a.x, s.a.y)$ and $(s.b.x, s.b.y)$, respectively, and ordered such that either $s.a.x < s.b.x$ or $s.a.x = s.b.x$ and $s.a.y < s.b.y$. That is, $s.b$ is never to the left of $s.a$, and if $s.a$ and $s.b$ have the same x coordinates, then $s.a$ is below $s.b$.

```
EQUALS(p, q)
    // tests whether p and q are the same point
1   if p.x == q.x and p.y == q.y
2       return TRUE
3   else return FALSE
```

```
ALGO-X(A)
1   for i = 1 to A.length
2       for j = 1 to A.length
3           if EQUALS(A[i].b, A[j].a)
4               for k = 1 to A.length
5                   if EQUALS(A[j].b, A[k].b) and EQUALS(A[i].a, A[k].a)
6                       return TRUE
7   return FALSE
```

Question 1: Analyze the asymptotic complexity of ALGO-X *(10')*

Question 2: Write an algorithm ALGO-Y that does exactly what ALGO-X does but with a better asymptotic complexity. Also, write the asymptotic complexity of ALGO-Y. *(20')*

►**Exercise 95.** Write an algorithm called TREE-TO-VINE that, given a binary search tree T, returns the same tree changed into a *vine*, that is, a tree containing exactly the same nodes but restructured so that no node has a left child (i.e., the returned tree looks like a linked list). The algorithm must not destroy or create nodes or use any additional memory other than what is already in the tree, and therefore must operate through a sequence of *rotations*. Write explicitly all the rotation algorithms used in TREE-TO-VINE. Also, analyze the complexity of TREE-TO-VINE. *(15')*

►**Exercise 96.** We say that a binary tree T is *perfectly balanced* if, for each node n in T, the number of keys in the left and right subtrees of n differ at most by 1. Write an algorithm called IS-PERFECTLY-BALANCED that, given a binary tree T returns TRUE if T is perfectly balanced, and FALSE otherwise. Also, analyze the complexity of IS-PERFECTLY-BALANCED. *(15')*

►**Exercise 97.** Two graphs G and H are *isomorphic* if there exists a *bijection* $f : V(G) \to V(H)$ between the vertexes of G and H (i.e., a one-to-one correspondence) such that any two vertices u and v in G are adjacent (in G) if and only if $f(u)$ and $f(v)$ are adjacent in H. The *graph-isomorphism* problem is the problem of deciding whether two given graphs are isomorphic.

Question 1: Is graph isomorphism in NP? If so, explain why and write a verification procedure. If not, argue why not. *(10')*

Question 2: Consider the following algorithm to solve the graph-isomorphism problem:

```
ISOMORPHIC(G, H)
1  if |V(G)| ≠ |V(H)|
2      return FALSE
3  A = V(G) sorted by degree    // A is a sequence of the vertices of G
4  B = V(H) sorted by degree    // B is a sequence of the vertices of H
5  for i = 1 to |V(G)|
6      if degree(A[i]) ≠ degree(B[i])
7          return FALSE
8  return TRUE
```

Is ISOMORPHIC correct? If so, explain at a high level what the algorithm does and informally but precisely why it works. If not, show a counter-example. *(10')*

►**Exercise 98.** Write an algorithm HEAP-PRINT-IN-ORDER(H) that takes a min heap H containing unique elements (no element appears twice in H) and prints the elements of H in increasing order. The algorithm must not modify H and may only use a constant amount of additional memory. Also, analyze the complexity of HEAP-PRINT-IN-ORDER. *(20')*

►**Exercise 99.** Write an algorithm BST-RANGE-WEIGHT(T, a, b) that takes a well balanced binary search tree T (or more specifically the root T of such a tree) and two keys a and b, with $a \leq b$, and returns the number of keys in T that are between a and b. Assuming there are $o(n)$ such keys, then the algorithm should have a complexity of $o(n)$, that is, strictly better than linear in the size of the tree. Analyze the complexity of BST-RANGE-WEIGHT. *(10')*

►**Exercise 100.** Let (a, b) represent an interval (or range) of values x such that $a \leq x \leq b$. Consider an array $X = \langle a_1, b_1, a_2, b_2, \ldots, a_n, b_n \rangle$ of $2n$ numbers representing n intervals (a_i, b_i), where $a_i = X[2i-1]$ and $b_i = X[2i]$ and $a_i \leq b_i$. Write an algorithm called SIMPLIFY-INTERVALS(X) that takes an array X representing n intervals, and simplifies X in-place. The "simplification" of a set of intervals X is a minimal set of intervals representing the *union* of all the intervals in X. Notice that the union of two disjoint intervals can not be simplified, but the union of two partially overlapping intervals can be simplified into a single interval. For example, a correct solution for the simplification of $X = \langle 3, 7, 1, 5, 10, 12, 6, 8 \rangle$ is $X = \langle 10, 12, 1, 8 \rangle$. An array X can be shrunk by setting its length (effectively removing elements at the end of the array). In this example, $X.length$ should be 4 after the execution of the simplification algorithm. Analyze the complexity of SIMPLIFY-INTERVALS. *(30')*

►**Exercise 101.** Write an algorithm SIMPLIFY-INTERVALS-FAST(X) that solves exercise 100 with a complexity of $O(n \log n)$. If your solution for exercise 100 already has an $O(n \log n)$ complexity, then simply say so. *(20')*

►**Exercise 102.** Consider the following algorithm:

```
ALGO-X(A, k)                    ALGO-Y(A, i)
1  i = 1                        1  while i < A.length
2  while i ≤ A.length           2      A[i] = A[i + 1]
3      if A[i] == k             3      i = i + 1
4          ALGO-Y(A, i)         4  A.length = A.length − 1 // discards last element
5      else i = i + 1
```

Analyze the complexity of ALGO-X and write an algorithm called BETTER-ALGO-X that does exactly the same thing, but with a strictly better asymptotic complexity. Analyze the complexity of BETTER-ALGO-X. *(20')*

►**Exercise 103.** Write an in-place partition algorithm called MODULO-PARTITION(A) that takes an array A of n numbers and changes A in such a way that (1) the final content of A is a permutation of the initial content of A, and (2) all the values that are equivalent to $0 \mod 10$ precede all the values equivalent to $1 \mod 10$, which precede all the values equivalent to $2 \mod 10$, etc. Being an in-place algorithm, MODULO-PARTITION must not allocate more than a constant amount of memory. For example, for an input array $A = \langle 7, 62, 5, 57, 12, 39, 5, 8, 16, 48 \rangle$, a correct result would be $A = \langle 12, 62, 5, 5, 16, 57, 7, 8, 48, 39 \rangle$. Analyze the complexity of MODULO-PARTITION. *(30')*

►**Exercise 104.** Write the *merge sort* algorithm and analyze its complexity. *(10')*

►**Exercise 105.** Write an algorithm called LONGEST-REPEATED-SUBSTRING(T) that takes a string T representing some text, and finds the longest string that occurs at least twice in T. The algorithm returns three numbers $begin_1, end_1$, and $begin_2$, where $begin_1 \leq end_1$ represent the first and last position of the *longest* substring of T that also occurs starting at another position $begin_2 \neq begin_1$ in T. If no such substring exist, then the algorithm returns "None." Analyze the time and space complexity of your algorithm. *(20')*

►**Exercise 106.** Answer the following questions on complexity theory. Recall that SAT is the Boolean satisfiability problem, which is a well-known NP-complete problem.

Question 1: A decision problem Q is polynomially-reducible to SAT. Can we say for sure that Q is NP-complete? *(2')*

Question 2: SAT is polynomially-reducible to a decision problem Q. Can we say for sure that Q is NP-complete? *(2')*

Question 3: A decision problem Q is polynomially reducible to a problem Q' and Q' is polynomially reducible to SAT. Can we say for sure that Q is in NP? *(2')*

Question 4: An algorithm A solves every instance of a decision problem Q of size n in $O(n^3)$ time. Also, Q is polynomially reducible to another problem Q'. Can we say for sure that Q' is in NP? *(2')*

Question 5: A decision problem Q is polynomially reducible to another decision problem Q', and an algorithm A solves Q' with complexity $O(n \log n)$. Can we say for sure that Q is in NP? *(2')*

Question 6: Consider the following decision problem Q: given a graph G, output 1 if G is connected (i.e., there exists a path between each pair of vertices) or 0 otherwise. Is Q in P? If so, outline an algorithm that proves it, if not argue why not. *(10')*

Question 7: Consider the following decision problem Q: given a graph G and an integer k, output 1 if G contains a cycle of size k. Is Q in NP? If so, outline an algorithm that proves it, if not argue why not. *(10')*

►**Exercise 107.** Consider an initially empty B-tree with minimum degree $t = 3$. Draw the B-tree after the insertion of the keys 84, 13, 36, 91, 98, 14, 81, 95, 12, 63, 31, and then after the additional insertion of the keys 65, 62, 187, 188, 57, 127, 6, 195, 25. *(10')*

►**Exercise 108.** Write an algorithm B-TREE-RANGE(T, k_1, k_2) that takes a B-tree T and two keys $k_1 \leq k_2$, and prints all the keys in T between k_1 and k_2 (inclusive). *(20')*

►**Exercise 109.** Write an algorithm called FIND-TRIANGLE(G) that takes a graph represented by its *adjacency list* G and returns true if G contains a triangle. A triangle in a graph G is a triple of vertices u, v, w such that all three edges (u, v), (v, w), and (u, w) are in G. Analyze the complexity of FIND-TRIANGLE. *(15')*

►**Exercise 110.** Write an algorithm MIN-HEAP-INSERT(H, k) that inserts a key k in a min-heap H. Also, illustrate the algorithm by writing the content of the array H after the insertion of keys 84, 13, 36, 91, 98, 14, 81, 95, 12, 63, 31, and then after the additional insertion of the key 15. *(15')*

►**Exercise 111.** Implement a priority queue by writing two algorithms:

- ENQUEUE(Q, x, p) enqueues an object x with priority p, and
- DEQUEUE(Q) extracts and returns an object from the queue.

The behavior of ENQUEUE and DEQUEUE is such that, if a call ENQUEUE(Q, x_1, p_1) is followed (not necessarily immediately) by another call ENQUEUE(Q, x_2, p_2), then x_1 is dequeued before x_2 unless $p_2 > p_1$. Implement ENQUEUE and DEQUEUE such that their complexity is $o(n)$ for a queue of n elements (i.e., strictly less than linear). *(20')*

►**Exercise 112.** Write an algorithm called MAX-HEAP-MERGE-NEW(H_1, H_2) that takes two max-heaps H_1 and H_2, and returns a new max-heap that contains all the elements of H_1 and H_2. MAX-HEAP-MERGE-NEW must create a *new* max heap, therefore it must allocate a new heap H and somehow copy all the elements from H_1 and H_2 into H without modifying H_1 and H_2. Also, analyze the complexity of MAX-HEAP-MERGE-NEW. *(20')*

►**Exercise 113.** Write an algorithm called BST-MERGE-INPLACE(T_1, T_2) that takes two binary-search trees T_1 and T_2, and returns a new binary-search tree by merging all the elements of T_1 and T_2. BST-MERGE-INPLACE is *in-place* in the sense that it must rearrange the nodes of T_1 and T_2 in a single binary-search tree without creating any new node. Also, analyze the complexity of BST-MERGE-INPLACE. *(20')*

►**Exercise 114.** Let A be an array of points in the 2D Euclidean space, each with its Cartesian coordinates $A[i].x$ and $A[i].y$. Write an algorithm MINIMUM-BOUNDING-RECTANGLE(A) that, given an array A of n points, in $O(n)$ time returns the smallest axis-aligned rectangle that contains all the points in A. MINIMUM-BOUNDING-RECTANGLE must return a pair of points corresponding to the bottom-left and top-right corners of the rectangle, respectively. *(10')*

►**Exercise 115.** Let A be an array of points in the 2D Euclidean space, each with its Cartesian coordinates $A[i].x$ and $A[i].y$. Write an algorithm LARGEST-CLUSTER(A, ℓ) that, given an array A of points and a length ℓ, returns the maximum number of points in A that are contained in a square of size ℓ. Also, analyze the complexity of LARGEST-CLUSTER. *(30')*

►**Exercise 116.** Consider the following algorithm that takes an array of numbers:

```
ALGO-X(A)
 1  i = 1
 2  j = 1
 3  m = 0
 4  c = 0
 5  while i ≤ |A|
 6      if A[i] == A[j]
 7          c = c + 1
 8      j = j + 1
 9      if j > |A|
10          if c > m
11              m = c
12          c = 0
13          i = i + 1
14          j = i
15  return m
```

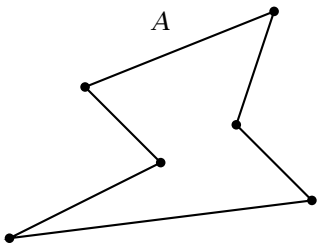
Question 1: Analyze the complexity of ALGO-X. *(5')*

Question 2: Write an algorithm that does exactly the same thing as ALGO-X but with a strictly better asymptotic time complexity. *(15')*

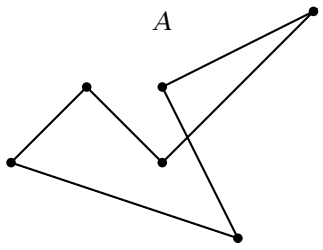
►**Exercise 117.** Write a THREE-WAY-MERGE(A, B, C) algorithm that merges three sorted sequences into a single sorted sequence, and use it to implement a THREE-WAY-MERGE-SORT(L) algorithm. Also, analyze the complexity of THREE-WAY-MERGE-SORT. *(20')*

►**Exercise 118.** Write an algorithm IS-SIMPLE-POLYGON(A) that takes a sequence A of 2D points, where each point $A[i]$ is defined by its Cartesian coordinates $A[i].x$ and $A[i].y$, and returns TRUE if A defines a simple polygon, or FALSE otherwise. Also, analyze the complexity of IS-SIMPLE-POLYGON. A polygon is *simple* if its line segments do not intersect.

Example:



IS-SIMPLE-POLYGON(A) = TRUE

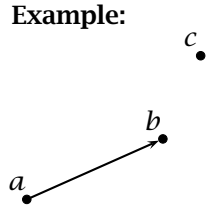


IS-SIMPLE-POLYGON(A) = FALSE

Hint: Use the following DIRECTION-ABC algorithm to determine whether a point c is *on the left side*, *collinear*, or *on the right side* of a segment ab:

```
DIRECTION-ABC(a, b, c)
1  d = (b.x − a.x)(c.y − a.y) − (b.y − a.y)(c.x − a.x)
2  if d > 0
3        return LEFT
4  elseif d == 0
5        return CO-LINEAR
6  else return RIGHT
```

Example:



DIRECTION-ABC(a, b, c) = LEFT

(20')

►**Exercise 119.** Implement a dictionary that supports *longest prefix matching*. Specifically, write the following algorithms:

- BUILD-DICTIONARY(W) takes a list W of n strings and builds the dictionary.
- LONGEST-PREFIX(k) takes a string k and returns the longest prefix of k found in the dictionary, or NULL if none exists. The time complexity of LONGEST-PREFIX(k) must be $o(n)$, that is, sublinear in the size n of the dictionary.

For example, assuming the dictionary was built with strings, "luna", "lunatic", "a", "al", "algo", "an", "anto", then if k is "algorithms", then LONGEST-PREFIX(k) should return "algo", or if k is "anarchy" then LONGEST-PREFIX(k) should return "an", or if k is "lugano" then LONGEST-PREFIX(k) should return NULL. *(20')*

►**Exercise 120.** Consider the following decision problem: given a set S of character strings, with characters of a fixed alphabet (e.g., the Roman alphabet), and given an integer k, return TRUE if there are at least k strings in S that have a common substring.

Question 1: Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. *(5')*

Question 2: Is the problem in P? Write an algorithm that proves it is, or argue the opposite. *(15')*

►**Exercise 121.** Draw a red-black tree containing the following set of keys, clearly indicating the color of each node.

$$\{8, 7, 7, 35, 23, 35, 13, 7, 23, 18, 3, 19, 22\}$$

(10')

►**Exercise 122.** Consider the following algorithm ALGO-X that takes an array A of n numbers:

```
ALGO-X(A)
1  return ALGO-XR(A, 0, 1, 2)
```

```
ALGO-XR(A, t, i, r)
1  while i ≤ A.length
2        if r == 0
3              if A[i] == t
4                    return TRUE
5        else if ALGO-XR(A, t − A[i], i + 1, r − 1)
6                    return TRUE
7        i = i + 1
8  return FALSE
```

Analyze the complexity of ALGO-X and then write an algorithm BETTER-ALGO-X that does exactly the same thing but with a strictly better time complexity. *(30')*

►**Exercise 123.** A Eulerian cycle in a graph is a cycle that goes through each edge exactly once. As it turns out, a graph contains a Eulerian cycle if (1) it is connected, and (2) all its vertexes have even degree. Write an algorithm EULERIAN(G) that takes a graph G represented as an adjacency matrix, and returns TRUE when G contains a Eulerian cycle. *(10')*

►**Exercise 124.** Consider a social network system that, for each user u, stores u's friends in a list *friends*(u). Implement an algorithm TOP-THREE-FRIENDS-OF-FRIENDS(u) that, given a user u, recommends the three other users that are not already among u's friends but are among the friends of most of u's friends. Also, analyze the complexity of the TOP-THREE-FRIENDS-OF-FRIENDS algorithm. *(20')*

►**Exercise 125.** Consider the following algorithm:

```
ALGO-X(A)
 1  for i = 3 to A.length
 2      for j = 2 to i − 1
 3          for k = 1 to j − 1
 4              x = A[i]
 5              y = A[j]
 6              z = A[k]
 7              if x > y
 8                  swap x ↔ y
 9              if y > z
10                  swap y ↔ z
11                  if x > y
12                      swap x ↔ y
13              if y − x == z − y
14                  return TRUE
15  return FALSE
```

Analyze the complexity of ALGO-X and write an algorithm called BETTER-ALGO-X(A) that does the same as ALGO-X(A) but with a strictly better asymptotic time complexity and with the same space complexity. *(20')*

►**Exercise 126.** The weather service stores the daily temperature measurements for each city as vectors of real numbers.

Question 1: Write an algorithm called HOT-DAYS(A, t) that takes an array A of daily temperature measurements for a city and a temperature t, and returns the maximum number of consecutive days with a recorded temperature above t. Also, analyze the complexity of HOT-DAYS(A, t). *(5')*

Question 2: Now imagine that a particular analysis would call the HOT-DAYS algorithm several times with the same series A of temperature measurements (but with different temperature values) and therefore it would be more efficient to somehow index or precompute the results. To do that, write the following two algorithms:

- A preprocessing algorithm called HOT-DAYS-INIT(A) that takes the series of temperature measurements A and creates an auxiliary data structure X (an index of some sort).
- An algorithm called HOT-DAYS-FAST(X, t) that takes the index X and a temperature t and returns the maximum number of consecutive days with a temperature above t. HOT-DAYS-FAST must run in *sub-linear time* in the size of A.

Also, analyze the complexity of HOT-DAYS-INIT and HOT-DAYS-FAST. *(25')*

►**Exercise 127.** Consider the following decision problem: given a sequence A of numbers and given an integer k, return TRUE if A contains either an increasing or a decreasing subsequence of length k. The elements of the subsequence must maintain their order in A but do not have to be contiguous.

Question 1: Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. *(10')*

Question 2: Is the problem in P? Write an algorithm that proves it is, or argue the opposite. *(20')*

►**Exercise 128.** Write an algorithm HEAP-DELETE(H, i) that, given a max-heap H, deletes the element at position i from H. *(10')*

►**Exercise 129.** Write an algorithm MAX-CLUSTER(A, d) that takes an array A of numbers (not necessarily integers) and a number d, and prints a maximal set of numbers in A that differ by at most d. The output can be given in any order. Your algorithm must have a complexity that is strictly better than $O(n^2)$. For example, with

$$A = \langle 7, 15, 16, 3, 10, 43, 8, 1, 29, 13, 4.5, 28 \rangle \quad d = 5$$

MAX-CLUSTER(A, d) would output $7, 3, 4.5, 8$ (or the same numbers in any other order) since those numbers differ by at most 5 and there is no larger set of numbers in A that differ by at most 5. Also, analyze the complexity of MAX-CLUSTER. *(20')*

►**Exercise 130.** Consider the following algorithm that takes a non-empty array of numbers

```
ALGO-X(A)
 1  B = make a copy of A
 2  i = 1
 3  while i ≤ B.length
 4        j = i + 1
 5        while j ≤ B.length
 6              if B[j] == B[i]
 7                    i = i + 1
 8                    swap B[i] ↔ B[j]
 9              j = j + 1
10        i = i + 1
11  q = B[1]
12  n = 1
13  m = 1
14  for i = 2 to B.length
15        if B[i] == q
16              n = n + 1
17              if n > m
18                    m = m + 1
19        else q = B[i]
20              n = 1
21  return m
```

Question 1: Briefly explain what ALGO-X does, and analyze the complexity of ALGO-X. *(10')*

Question 2: Write an algorithm called BETTER-ALGO-X that is functionally identical to ALGO-X but with a strictly better complexity. Analyze the complexity of BETTER-ALGO-X. *(10')*

►**Exercise 131.** Write the *heap-sort* algorithm and then illustrate how *heap-sort* processes the following array in-place:

$$A = \langle 33, 28, 23, 48, 32, 46, 40, 12, 21, 41, 14, 37, 38, 0, 25 \rangle$$

In particular, show the content of the array at each main iteration of the algorithm. *(20')*

►**Exercise 132.** Write an algorithm BST-PRINT-LONGEST-PATH(T) that, given a binary search tree T, outputs the sequence of nodes (values) of the path from the root to any node of maximal depth. Also, analyze the complexity of BST-PRINT-LONGEST-PATH. *(30')*

►**Exercise 133.** Consider insertion in a binary search tree.

Question 1: Write a valid insertion algorithm BST-INSERT. *(10')*

Question 2: Illustrate how BST-INSERT works by drawing the binary search tree resulting from the insertion of the following keys in this order:

$$33, 28, 23, 48, 32, 46, 40, 12, 21, 41, 14, 37, 38, 0, 25$$

Also, if the resulting tree is not already of minimal depth, write an alternative insertion order that would result in a tree of minimal depth. *(10')*

Question 3: Write an algorithm BEST-BST-INSERT-ORDER(A) that takes an array of numbers A and outputs the elements of A in an order that, if used with BST-INSERT would lead to a binary search tree of minimal depth. *(10')*

►**Exercise 134.** Write an algorithm called FIND-NEGATIVE-CYCLE that, given a weighted directed graph $G = (V, E)$, with weight function $w : E \to \mathbb{R}$, finds and outputs a negative-weight cycle in G if one such cycle exists. Also, analyze the complexity of FIND-NEGATIVE-CYCLE. *(20')*

►**Exercise 135.** Consider a text composed of n lines of up to 80 characters each. The text is stored in an array T where each line $T[i]$ is an array of characters containing words separated by a single space.

Question 1: Write an algorithm SORT-LINES-BY-WORD-COUNT(T) that, with a worst-case complexity of $O(n)$, sorts the lines in T in non-decreasing order of the number of words in the line. (**Hint:** lines have at most 80 characters, so the number of words in a line is also limited.) *(20')*

Question 2: If you did not already do that for exercise 1, write an *in-place* variant of the SORT-LINES-BY-WORD-COUNT algorithm. This algorithm, called SORT-LINES-BY-WORD-COUNT-IN-PLACE, must also have a $O(n)$ complexity to sort the set of lines, and may use only a constant amount of extra space to do that. *(20')*

►**Exercise 136.** Consider a weighted undirected graph $G = (V, E)$ representing a group of programmers and their affinity for team work, such that the weight $w(e)$ of an edge $e = (u, v)$ is a number representing the ability of programmers u and v to work together on the same project. Write an algorithm BEST-TEAM-OF-THREE that outputs the best team of three programmers. The value of a team is considered to be the lowest affinity level between any two members of the team. So, the best team is the group of programmers for which the lowest affinity level between members of the group is maximal. *(20')*

►**Exercise 137.** Write an algorithm MAXIMAL-NON-ADJACENT-SEQUENCE-WEIGHT(A) that, given a sequence of numbers $A = \langle a_1, a_2, ..., a_n \rangle$, computes, with worst-case complexity $O(n)$, the maximal weight of any sub-sequence of non-adjacent elements in A. A sub-sequence of non-adjacent elements may include a_i or a_{i+1} but not both, for all i. For example, with $A = \langle 2, 9, 6, 2, 6, 8, 5 \rangle$, MAXIMAL-NON-ADJACENT-SEQUENCE-WEIGHT(A) should return 20. (**Hint:** use a dynamic programming algorithm that scans the input sequence once.) *(20')*

►**Exercise 138.** Consider a trie rooted at node T that represents a set of character strings. For simplicity, assume that characters are from the Roman alphabet and that the letters of the alphabet are encoded with numeric values between 1 and 26. Write an algorithm PRINT-TRIE(T) that prints all the strings stored in the trie. *(20')*

►**Exercise 139.** Write an algorithm PRINT-IN-THREE-COLUMNS(A) that takes an array of words A and prints all the words in A, in the given order left-to-right and top-to-bottom, such that the words are left-aligned in three columns. Words must be separated by at least one space horizontally, but in order to align words, the algorithm might have to print more spaces between words. For example, if A contains the words *exam, algorithms, asymptotic, complexity, graph, greedy, lugano, np, quicksort, retake, september,* then the output should be

```
exam       algorithms asymptotic
complexity graph      greedy
lugano     np         quicksort
retake     september
```

(20')

►**Exercise 140.** Consider a binary search tree.

Question 1: Write an algorithm BST-MEDIAN(T) that takes the root T of a binary search tree and returns the median element contained in the tree. Also analyze the complexity of BST-MEDIAN(T). Can you do better? *(10')*

Question 2: Assume now that the tree is balanced and also that each node t has an attribute $t.weight$ corresponding to the total number of nodes in the subtree rooted at t (including t itself). Write an algorithm BETTER-BST-MEDIAN(T) that improves on the complexity of BST-MEDIAN. Analyze the complexity of BETTER-BST-MEDIAN. *(10')*

►**Exercise 141.** Consider the following decision problem. Given a set of strings S, a number w, and a number k, output *YES* when there are at least k strings in S that share a common sub-string of length w, or *NO* otherwise. For example, if S contains the strings *exam, algorithms, asymptotic, complexity, graph, greedy, lugano, np, quicksort, retake, september, theory, practice, programming, math, art, truth, justice,* with $w = 2$ and $k = 3$ the output should be *YES*, because the 3 strings *graph, greedy,* and *programming* share a common substring "gr" of length 2. The output should also be *YES* for $w = 3$ and $k = 3$ and for $w = 2$ and $k = 4$, but it should be *NO* for $w = 3$ and $k = 4$.

Question 1: Is this problem in NP? Write an algorithm that proves it is, or argue that it is not. *(10')*

Question 2: Is this problem in P? Write an algorithm that proves it is, or argue that it is not. (**Hint:** a string of length ℓ has $O(\ell^2)$ sub-strings of any length.) *(20')*

►**Exercise 142.** Consider the following sorting problem: you must reorder the elements of an array of numbers in-place so that odd numbers are in odd positions while even numbers are in even positions. If there are more even elements than odd ones in A (or vice-versa) then those additional elements will be grouped at the end of the array. For example, with an initial sequence

$$A = \langle 50, 47, 92, 78, 76, 7, 60, 36, 59, 30, 50, 43 \rangle$$

the result could be this:

$$A = \langle 47, 50, 7, 78, 59, 76, 43, 92, 36, 60, 30, 50 \rangle$$

Question 1: Write an algorithm called ALTERNATE-EVEN-ODD(A) that sorts A in place as explained above. Also, analyze the complexity of ALTERNATE-EVEN-ODD. (You might want to consider question 2 before you start solving this problem.) *(20')*

Question 2: If you have not done so already, write a variant of ALTERNATE-EVEN-ODD that runs in $O(n)$ steps for an array A of n elements. *(10')*

►**Exercise 143.** Write an algorithm called FOUR-CYCLE(G) that takes a directed graph represented with its adjacency matrix G, and that returns *true* if and only if G contains a 4-cycle. A 4-cycle is a sequence of four distinct vertexes a, b, c, d such that there is an arc from a to b, from b to c, from c to d, and from d to a. Also, analyze the complexity of FOUR-CYCLE(G). *(20')*

►**Exercise 144.** Write an algorithm FIND-EQUAL-DISTANCE(A) that takes an array A of numbers, and returns four distinct elements a, b, c, d of A such that $a - b = c - d$, or NIL if no such elements exist. FIND-EQUAL-DISTANCE must run in $O(n^2 \log n)$ time. *(20')*

►**Exercise 145.** Consider the following algorithm that takes an array of numbers:

```
ALGO-X(A)
 1  i = 1
 2  while i < A.length
 3       if A[i] > A[i + 1]
 4            swap A[i] ↔ A[i + 1]
 5       p = i
 6       q = i + 1
 7       for j = i + 2 to A.length
 8            if A[j] < A[p]
 9                 p = j
10            else if A[j] > A[q]
11                      q = j
12       swap A[i] ↔ A[p]
13       swap A[i + 1] ↔ A[q]
14       i = i + 2
```

Question 1: Explain what ALGO-X does and analyze its complexity. *(5')*

Question 2: Write an algorithm BETTER-ALGO-X that is functionally equivalent to ALGO-X but with a strictly better time complexity. *(15')*

►**Exercise 146.** Consider the following definition of the height of a node t in a binary tree:

$$height(t) = \begin{cases} 0 & \text{if } t == \text{NIL} \\ 1 + \max\{height(t.left), height(t.right)\} & \text{otherwise.} \end{cases}$$

Question 1: Write an algorithm HEIGHT(t) that computes the height of a node t. Also, analyze the complexity of your HEIGHT algorithm when t is the root of a tree of n nodes. *(5')*

Question 2: Consider now a binary search tree in which each node t has an attribute $t.height$ that denotes the height of that node. Write a constant-time rotation algorithm LEFT-ROTATE(t) that performs a left rotation around node t and also updates the *height* attributes as needed. *(5')*

►**Exercise 147.** Consider the following classic insertion algorithm for a binary search tree:

```
BST-INSERT(t, k)
1  if t == NIL
2       return NEW-NODE(k)
3  else if k ≤ t.key
4            t.left = BST-INSERT(t.left, k)
5       else t.right = BST-INSERT(t.right, k)
6  return t
```

Write an algorithm SORT-FOR-BALANCED-BST(A) that takes an array of numbers A, and prints the elements of A so that, if passed to BST-INSERT, the resulting BST would be of minimal height. Also, analyze the complexity of your solution. *(20')*

►**Exercise 148.** Consider the array of numbers:

$$A = \langle 69, 36, 68, 18, 36, 36, 50, 9, 36, 36, 18, 18, 8, 10 \rangle$$

Question 1: Does A satisfy the *max-heap* property? If not, fix it by swapping two elements. *(5')*

Question 2: Write an algorithm MAX-HEAP-INSERT(H, k) that inserts a key k in a max-heap H. *(10')*

Question 3: Illustrate the behavior of MAX-HEAP-INSERT by applying it to array A (possibly corrected). In particular, write the content of the array after the insertion of each of the following keys, in this order: $69, 50, 60, 70$. *(5')*

►**Exercise 149.** Consider the following algorithm that takes an array of numbers:

```
ALGO-Y(A)
 1  a = 0
 2  for i = 1 to A.length − 1
 3       for j = i + 1 to A.length
 4            x = 0
 5            for k = i to j
 6                 if A[k] is even:
 7                      x = x + 1
 8                 else x = x − 1
 9            if x == 0 and j − i > a
10                 a = j − i
11 return a
```

Question 1: Explain what ALGO-Y does and analyze its complexity. *(5')*

Question 2: Write an algorithm BETTER-ALGO-Y that is functionally equivalent to ALGO-Y but with a strictly better time complexity. Also analyze the time complexity of BETTER-ALGO-Y. *(10')*

Question 3: If you have not already done so for question 2, write a BETTER-ALGO-Y that is functionally equivalent to ALGO-Y but that runs in time $O(n)$. *(15')*

►**Exercise 150.** Write an algorithm THREE-WAY-PARTITION(A, v) that takes an array A of n numbers, and partitions A *in-place* in three parts, some of which might be empty, so that the left part $A[1 \ldots p-1]$ contains all the elements less than v, the middle part $A[p \ldots q-1]$ contains all the elements equal to v, and the right part $A[q \ldots n]$ contains all the elements greater than v. THREE-WAY-PARTITION must return the positions p and q and must run in time $O(n)$. *(20')*

►**Exercise 151.** A DNA strand is a sequence of nucleotides, and can be represented as a string over the alphabet $\Sigma = \{\mathsf{A}, \mathsf{C}, \mathsf{G}, \mathsf{T}\}$. Consider the problem of determining whether two DNA strands s_1 and s_2 are k-related in the sense that they share a sub-sequence of at least k nucleotides.

Question 1: Is the problem in NP? Write an algorithm that proves it is, or argue that it is not. *(10')*

Question 2: Is the problem in P? Write an algorithm that proves it is, or argue that it is not. *(20')*

►**Exercise 152.** Consider the following algorithm that takes an array of numbers:

```
ALGO-X(A)
 1  y = −∞
 2  i = 1
 3  j = 1
 4  x = 0
 5  while i ≤ A.length
 6        x = x + A[j]
 7        if x > y
 8              y = x
 9        if j == A.length
10              i = i + 1
11              j = i
12              x = 0
13        else j = j + 1
14  return y
```

Question 1: Explain what ALGO-X does and analyze its complexity. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X that is functionally equivalent to ALGO-X but with a strictly better time complexity. *(20')*

►**Exercise 153.** Write an algorithm MAXIMAL-CONNECTED-SUBGRAPH(G) that takes an undirected graph $G = (V, E)$ and prints the vertices of a maximal connected subgraph of G. *(20')*

►**Exercise 154.** A system collects the positions of cars along a highway that connects two cities, A and B. The positions are grouped by direction in two arrays, A and B. Thus A contains the distances in kilometers from city A of the cars traveling towards city A. Write an algorithm CONGESTION(A) that takes the array A and prints a list of congested sections of the highway. A congested interval is a contiguous stretch of highway of 1km or more in which the density of cars is more than 50 cars per kilometer. CONGESTION(A) must run in $O(n \log n)$ time. *(20')*

►**Exercise 155.** The following matrix represents a directed graph over vertices $a, b, c, \ldots, \ell$. Rows and columns represent the source and destination of edges, respectively.

	a	b	c	d	e	f	g	h	i	j	k	ℓ
a			1									
b	1		1								1	
c				1	1							
d	1						1	1				
e				1		1					1	
f				1			1					
g												
h									1	1		
i												
j								1	1			
k												1
ℓ											1	

Write the graph and the *DFS numbering* of the vertexes using the DFS algorithm. Every iteration through vertexes or adjacent edges is performed in alphabetic order. (*Hint:* the DFS numbering of a vertex v is a pair of numbers representing the "time" at which DFS discovers v and the time DFS leaves v.) *(20')*

►**Exercise 156.** Consider an array A of n numbers that is initially sorted, in ascending order, and then modified so that k of its elements are decreased in value.

Question 1: Write an algorithm that sorts A *in-place* in time $O(kn)$. *(10')*

Question 2: Write an algorithm that sorts A in time $O(n + k \log k)$ but not necessarily in-place. *(20')*

►**Exercise 157.** Consider the decision version of the well-known *vertex cover* problem: given a graph $G = (V, E)$ and an integer k, output 1 if G contains a vertex cover of size k. A vertex cover is a set of vertexes $S \subseteq V$ such that, for each edge $(u, v) \in E$, either vertex u is in S or vertex v is in S. Write an algorithm that proves that vertex cover is in NP. *(20')*

►**Exercise 158.** Write an algorithm that transforms a min-heap H into a max-heap *in-place*. *(10')*

►**Exercise 159.** We say that two words x and y are *linked* to each other if they differ by a single letter, or more specifically by one edit operation, meaning an insertion, a deletion, or a change in a single character. For example, "fun" and "pun" are linked, as are "flower" and "lower", "port" and"post", "canton" and "cannon", and "cat" and "cast".

Question 1: Write an algorithm LINKED(x, y) that takes two words x and y and, in lienar time, returns TRUE if x and y are linked to each other, or FALSE otherwise. *(10')*

Question 2: Write an algorithm WORD-CHAIN(W, x, y) that takes an array of words W and two words x and y, and outputs a minimal sequence of words $x, w_1, w_2, \ldots, y$ that starts with x and ends with y where $w_1, w_2, \ldots$ are all words from W, and each word in the sequence is linked to the words adjacent to it. For example, if W is a dictionary of English words, and x and y are "first" and "last", respectively, then the output might be: *first fist list last*. *(30')*

►**Exercise 160.** Write an algorithm MAX-HEAP-INSERT(H, k) that inserts a new value k in a max-heap H. Briefly analyze the complexity of your solution. *(10')*

►**Exercise 161.** Consider an algorithm FIND-ELEMENTS-AT-DISTANCE(A, k) that takes an array A of n integers sorted in non decreasing order and returns TRUE if and only if A contains two elements a_i and a_j such that $a_i - a_j = k$.

Question 1: Write a version of the FIND-ELEMENTS-AT-DISTANCE algorithm that runs in $O(n \log n)$ time. Briefly analyze the complexity of your solution. *(10')*

Question 2: Write a version of the FIND-ELEMENTS-AT-DISTANCE algorithm that runs in $O(n)$ time. Briefly analyze the complexity of your solution. *(20')*

►**Exercise 162.** Write an algorithm PARTITION-PRIMES-COMPOSITES(A) that takes an array A of n integers such that $1 < A[i] \leq m$ for all i, and partitions A in-place so that all primes precede all composites in A. Analyze the complexity of your solution as a function of n and m. Recall that an integer greater than 1 is *prime* if it is divisible by only two positive integers (itself and 1) or otherwise it is *composite*. *(20')*

►**Exercise 163.** Consider the following classic insertion algorithm for a binary search tree:

```
BST-INSERT(t, k)
1  if t == NIL
2        return NEW-NODE(k)
3  else if k ≤ t.key
4              t.left = BST-INSERT(t.left, k)
5        else t.right = BST-INSERT(t.right, k)
6  return t
```

Write an algorithm SORT-FOR-BALANCED-BST(A) that takes an array of numbers A, and prints the elements of A in a new order so that, if the printed sequence is passed to BST-INSERT, the resulting BST would be of minimal height. Also, analyze the complexity of your solution. *(20')*

►**Exercise 164.** Consider a game in which, given a multiset of positive numbers A (possibly with repeated values) a player can simplify A by removing, one at a time, an element a_k if there are two other elements a_i, a_j such that $a_i + a_j = a_k$.

Question 1: Write an algorithm called MINIMAL-SIMPLIFIED-SUBSET(A) that, given a multiset A of n numbers, returns a minimal simplified subset $X \subseteq A$. The result X is *minimal* in the sense that no smaller set can be obtained with a sequence of simplifications starting from A. For example, with $A = \{7, 89, 11, 88, 106, 4, 28, 71, 17\}$, a valid result would be $X = \{7, 89, 4, 71, 17\}$. Briefly analyze the complexity of your solution. *(10')*

Question 2: Write a MINIMAL-SIMPLIFIED-SUBSET(A) algorithm that runs in $O(n^2)$. If you have already done so for exercise 1, then simply say so. *(20')*

►**Exercise 165.** Consider the following algorithm that takes an integer n as input:

```
ALGORITHM-X(n)
1  c = 0
2  a = n
3  while a > 1
4       b = 1
5       while b ≤ a^2
6            c = c + 1
7            b = 2b
8       a = a/2
9  return c
```

Write the complexity of ALGORITHM-X as a function of n. Justify your answer. *(10')*

►**Exercise 166.** Write an algorithm FIND-CYCLE(G) that, given a directed graph G, returns TRUE if and only if G contains a cycle. You may assume the representation of your choice for G. *(20')*

►**Exercise 167.** A breadth-first search over a graph G returns a vector π that represents the resulting breadth-first tree, where the parent $\pi[v]$ of a vertex v is the next-hop from v on the tree towards the source of the breadth-first search.

Question 1: Write an algorithm BFS-FIRST-COMMON-ANCESTOR(π, u, v) that finds the first common ancestor of two given nodes in the breadth-first tree, or NULL if u and v are not connected in G. The complexity of BFS-FIRST-COMMON-ANCESTOR must be $O(n)$. Briefly analyze the space complexity of your solution. *(10')*

Question 2: Write an algorithm BFS-FIRST-COMMON-ANCESTOR-2(π, D, u, v) that is also given the distance vector D resulting from the same breadth first search. BFS-FIRST-COMMON-ANCESTOR-2 must be functionally equivalent to BFS-FIRST-COMMON-ANCESTOR (as defined in Exercise 1) but with space complexity $O(1)$. *(20')*

►**Exercise 168.** Consider the height and the black height of a red-black tree.

Question 1: What are the minimum and maximum heights of a red-black tree containing 10 keys? Exemplify your answers by drawing a minimal and a maximal tree. Clearly identify each node as red or black. *(10')*

Question 2: What are the minimum and maximum *black* heights of a red-black tree containing 10 keys? Exemplify your answers by drawing a minimal and a maximal tree. Clearly identify each node as red or black. *(10')*

►**Exercise 169.** Consider an algorithm BST-FIND-SUM(T, v) that, given a binary search tree T containing n distinct numeric keys, and given a target value v, finds and returns two nodes in T whose keys add up to v. The algorithm returns NULL if no such keys exist in T. BST-FIND-SUM may not modify the tree, and may only use a constant amount of memory.

Question 1: Write BST-FIND-SUM. You may use the basic algorithms that operate on binary search trees (BST-MIN, BST-SUCCESSOR, BST-SEARCH, etc.) without defining them explicitly. *(10')*

Question 2: Write a variant of BST-FIND-SUM(T, v) that works in $O(n)$ time. If your solution to Exercise 1 already has this complexity bound, then simply say so. *(20')*

►**Exercise 170.** Consider this decision problem: given a set of integers $X = \{x_1, x_2, \ldots, x_n\}$, and an integer k, return 1 if there are k elements in X that are pairwise relatively prime, or return 0 otherwise. Two integers are relatively prime if their only common divisor is 1. For example, for $X = \{5, 6, 10, 14, 18, 21, 49\}$ and $k = 3$, the result is 1, since the 3 elements $5, 18, 49$ are pairwise relatively prime (5 and 18 have no common divisor other than 1, and the same is true for 5 and 49, and 18 and 49). However, for the same set $X = \{5, 6, 10, 14, 18, 21, 49\}$ and $k = 4$, the solution is 0, since no four elements from X are all pairwise relatively prime.

Question 1: Is this problem in NP? Write an algorithm that proves it is, or argue that it is not. *(20')*

Question 2: **(BONUS)** Is this problem NP-hard? Prove it. *(60')*

►**Exercise 171.** You are given a square matrix $M \in \mathbf{R}^{n \times n}$ whose elements are sorted both row-wise and column-wise. In other words, rows and columns are non-decreasing sequences. Formally, for every element $m_{i,j} \in M$, $(j < n \Rightarrow m_{i,j} \leq m_{i,j+1}) \wedge (i < n \Rightarrow m_{i,j} \leq m_{i+1,j})$. Write an algorithm SEARCH-IN-SORTED-MATRIX(M, x) that returns TRUE if $x \in M$ or FALSE otherwise. The time complexity of SEARCH-IN-SORTED-MATRIX must be $O(n \log n)$. Justify that your solution has such a complexity. *(20')*

►**Exercise 172.** Consider the following algorithm that takes an array A of positive integers:

```
ALGO-X(A)
 1  B = copy of A
 2  i = 1
 3  x = 1
 4  while i ≤ A.length
 5       B[i] = B[i] − 1
 6       if B[i] == 0
 7            B[i] = A[i]
 8            i = i + 1
 9       else x = x + 1
10            i = 1
11  return x
```

Question 1: Briefly explain what ALGO-X does and analyze the complexity of ALGO-X. *(10')*

Question 2: Write an algorithm called BETTER-ALGO-X that is functionally identical to ALGO-X but with a strictly better complexity. Analyze the complexity of BETTER-ALGO-X. *(10')*

►**Exercise 173.** Consider the following algorithm that takes an array A of numbers:

```
ALGO-Y(A)
 1  i = 2
 2  j = 1
 3  x = −∞
 4  while i ≤ A.length
 5       if |A[i] − A[j]| > x
 6            x = |A[i] − A[j]|
 7       j = j + 1
 8       if j == i
 9            i = i + 1
10            j = 1
11  return x
```

Question 1: Briefly explain what ALGO-Y does and analyze the complexity of ALGO-Y. *(10')*

Question 2: Write an algorithm called BETTER-ALGO-Y that is functionally identical to ALGO-Y but with a complexity $O(n)$. *(10')*

►**Exercise 174.** Write an algorithm BTREE-LOWER-BOUND(T, k) that, given a B-tree T and a value k, returns the least key v in T such that $k \leq v$, or NULL if no such key exist. Also, analyze the complexity of BTREE-LOWER-BOUND. Recall that a node x in a B-tree has the following properties:

$x.n$ is the number of keys, $X.key[1] \leq x.key[2] \leq \ldots x.key[x.n]$ are the keys, $x.leaf$ tells whether x is a leaf, and $x.c[1], x.c[2], \ldots, x.c[x.n+1]$ are the pointers to x's children. *(20')*

►**Exercise 175.** Write an algorithm BST-LEAST-DIFFERENCE(T) that, given a binary search tree T containing numeric keys, returns in $O(n)$ time the minimal distance between any two keys in the tree. *(20')*

►**Exercise 176.** A connected component of an undirected graph G is a maximal set of vertices that are connected to each other (directly or indirectly). Thus the vertices of a graph can be partitioned into connected components. Write an algorithm CONNECTED-COMPONENTS(G) that, given an undirected graph G, returns the number of connected components in G. Also, analyze the complexity of CONNECTED-COMPONENTS. *(20')*

►**Exercise 177.** Rank the following functions in decreasing order of growth by indicating their rank next to the function, as in the first line (n^{n^n} is the fastest growing function). If any two functions f_i and f_j are such that $f_i = \Theta(f_j)$, then rank them at the same level. *(10')*

function	*rank*
$f_0(n) = n^{n^n}$	1
$f_1(n) = \log^2(n)$	
$f_2(n) = n!$	
$f_3(n) = \log(n^2)$	
$f_4(n) = n$	
$f_5(n) = \log(n!)$	
$f_6(n) = \log\log n$	
$f_7(n) = n \log n$	
$f_8(n) = \sqrt{n^3}$	
$f_9(n) = 2^n$	

Hint: as a reminder, consider the following mathematical definitions and facts: (definition of factorial) $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$; (facts about the logarithm) $\log(ab) = \log a + \log b$, and therefore $\log(a^k) = k \log a$.

►**Exercise 178.** Write an algorithm called MINIMAL-COVERING-SQUARE(P) that takes a sequence P of n points in the 2D Euclidean plane, each defined by its Cartesian coordinates $P[i].x$ and $P[i].y$, and returns the area of a minimal axis-aligned square that covers all points in P. An axis-aligned square is one in which the sides are parallel to X and Y axes. MINIMAL-COVERING-SQUARE must run in time $O(n)$. *(10')*

►**Exercise 179.** A sequence of numbers is called *unimodal* if it is first strictly increasing and then strictly decreasing. For example, the sequence $1, 5, 19, 17, 12, 8, 5, 3, 2$ is unimodal, while the sequence $1, 5, 3, 7, 4, 2$ is not. Write an algorithm UNIMODAL-FIND-MAXIMUM(A) that finds the maximum of a unimodal sequence A of n numbers in time $O(\log n)$. *(20')*

►**Exercise 180.** Consider the following algorithm ALGO-X(A, k) that takes an array A of n objects and an integer k:

```
ALGO-X(A, k)
1  l = −∞
2  r = +∞
3  for i = 1 to A.length − k
4       for j = i + 1 to A.length
5            if ALGO-Y(A, i, j) ≥ k
6                 if r − l > j − i
7                      l = i
8                      r = j
9  return l, r
```

```
ALGO-Y(A, a, b)
1  m = 1
2  for i = a to b
3       c = 1
4       for j = i + 1 to b
5            if A[i] == A[j]
6                 c = c + 1
7       if c > m
8            m = c
9  return m
```

Question 1: Explain what ALGO-X(A, k) does and analyze its complexity. Do not simply paraphrase the code. Instead, explain the high level semantics, independent of the code. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X(A, k) with the same functionality as ALGO-X(A, k), but with a strictly better complexity. Also, analyze the complexity of BETTER-ALGO-X(A, k). *(20')*

►**Exercise 181.** An algorithm THREE-WAY-PARTITION$(A, \mathit{begin}, \mathit{end})$ chooses a pivot element from the sub-array $A[\mathit{begin} \ldots \mathit{end} - 1]$, and partitions that sub-array in-place into three parts (two of which might be empty): $A[\mathit{begin} \ldots q_1 - 1]$ containing all the elements less than the pivot, $A[q_1 \ldots q_2 - 1]$ containing all the elements equal to the pivot, and $A[q_2 \ldots \mathit{end} - 1]$ containing all elements greater than the pivot.

Question 1: Write a THREE-WAY-PARTITION$(A, \mathit{begin}, \mathit{end})$ algorithm that runs in time $O(n)$, where $n = \mathit{end} - \mathit{begin}$, and that returns the partition boundaries q_1, q_2. You may assume that $\mathit{begin} < \mathit{end}$. *(20')*

Question 2: Use the THREE-WAY-PARTITION algorithm to write a better variant of the classic quick-sort algorithm. Also, describe in which cases this variant would perform significantly better than the classic algorithm. *(10')*

►**Exercise 182.** the following algorithm SUM(A, s) takes an array A of n numbers and a number s. Describe what SUM(A, s) does at a high level and analyze its complexity in the best and worst cases. Justify your answer by clearly describing the best- and worst-case input, as well as the behavior of the algorithm in each case. *(20')*

```
SUM(A, s)
1  return SUM-R(A, s, 1, A.length)
```

```
SUM-R(A, s, b, e)
1  if b > e and s == 0
2       return TRUE
3  elseif b ≤ e and SUM-R(A, s, b + 1, e)
4       return TRUE
5  elseif b ≤ e and SUM-R(A, s − A[b], b + 1, e)
6       return TRUE
7  else return FALSE
```

►**Exercise 183.** Big Brother tracks a set of m cell-phone users by recording every cell antenna the user connects to. In particular, for each user u_i, Big Brother stores a time-ordered sequence $S_i = (t_1, a_1), (t_2, a_2), \ldots$ that records that user u_i was connected to antenna a_1 starting at time t_1, and later switched to antenna a_2 at time $t_2 > t_1$, and so on. Write an algorithm called GROUP-OF-K$(S_1, S_2, \ldots, S_m, k)$ that finds whether there is a time t^* when a group of at least k users are connected to the same antenna. In this case, GROUP-OF-K must output the time t^* and the antenna a^*. Otherwise, GROUP-OF-K must output NULL. GROUP-OF-K must run in time $O(n \log m)$ where n is the total number of entries in all the sequences, so $n = |S_1| + |S_2| + \cdots + |S_m|$. You may use common data structures and algorithms without specifying those algorithms completely. *(20')*

►**Exercise 184.** Consider the following algorithm that takes an array A of integers:

```
ALGO-X(A)
 1  i = 1
 2  j = A.length + 1
 3  while i < j
 4      if A[i] ≡ 0 mod 2
        // A[i] is even
 5          j = j − 1
 6          v = A[i]
 7          ALGO-Y(A, i, j)
 8          A[j] = v
 9      else i = i + 1
10  return j
```

```
ALGO-Y(A, p, q)
1  while p < q
2      A[p] = A[p + 1]
3      p = p + 1
```

Question 1: Briefly explain what ALGO-X does and analyze the complexity of ALGO-X. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X that is functionally identical to ALGO-X but with a strictly better complexity. Also briefly analyze the complexity of BETTER-ALGO-X. *(10')*

►**Exercise 185.** Write an algorithm BTREE-PRINT-RANGE(T, a, b) that, given a B-tree T and two values $a < b$, prints all the keys k in T that are between a and b, that is, $a < k < b$. Recall that a node x in a B-tree has the following properties: $x.n$ is the number of keys, $x.key[1] \le x.key[2] \le \ldots x.key[x.n]$ are the keys, $x.leaf$ tells whether x is a leaf, and $x.c[1], x.c[2], \ldots, x.c[x.n+1]$ are the pointers to x's children. *(20')*

►**Exercise 186.** Consider the following decision problem: given a weighted graph G and a number k, where $w(e)$ is the weight of an edge $e = (u, v) \in E(G)$, return TRUE if and only if there are at least two nodes u and v at distance $d(u, v) = k$. Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. Is the problem in P? Write an algorithm that proves it is, or argue the opposite. Recall that the distance $d(u, v)$ in a graph is the minimal length of any path connecting u and v. *(20')*

►**Exercise 187.** A highway traffic app sends the coordinates of each vehicle to a central server that reports on congested sections of highway. For simplicity, consider the highway as a straight line in which each position is identified by a single x coordinate. Write an algorithm MOST-CONGESTED-SEGMENT(A, ℓ) that, given an array A of vehicle positions and a length ℓ, outputs the position of a maximally congested highway segment of length at most ℓ. A segment of highway between positions x and $x + \ell$ is considered maximally congested if there are no other segments of length at most ℓ with more vehicles. Coordinates as well as the length ℓ are real numbers, not necessarily integers; ℓ is positive (it is a distance). *(20')*

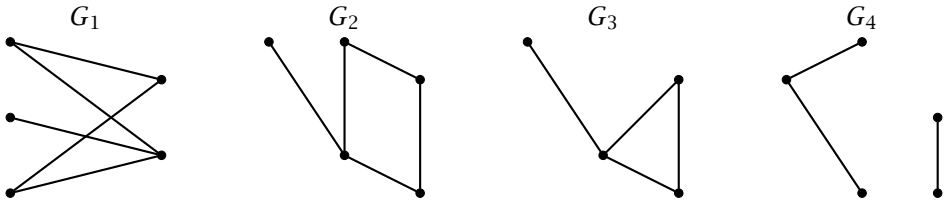
►**Exercise 188.** Consider the following decision problem: given a graph G represented as an adjacency matrix G, and an integer k, return TRUE if and only if there are at least k nodes $v_1, v_2, \ldots, v_k$ in G that form a fully connected sub-graph of G, meaning that for every pair $i, j \in 1, \ldots, k$, edge (v_i, v_j) is in G. Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. *(20')*

►**Exercise 189.** Write an algorithm MAX-HEAP-TOP-THREE(H) that takes a heap H and prints the three highest values stored in the heap. The algorithm must run in $O(1)$ time, may not allocate more than a constant amount of memory, and may not modify the heap in any way. If the heap contains less than three values, then MAX-HEAP-TOP-THREE must print whatever elements exist. *(20')*

►**Exercise 190.** Let P be a sequence of points representing an alpine road, where each point $p \in P$ is defined by two coordinates $p.x$ and $p.y$ where $p.x$ is the distance from the beginning of the road and $p.y$ is the elevation (meters above see level). Write an algorithm LONGEST-STRETCH(P, h) that takes a sequence of points P and an altitude range (difference) h, and returns the maximal length of a stretch of road that remains within an altitude range of at most h. For example, if $h = 0$, the algorithm must return the maximal length of road that is absolutely flat (that is, contiguous points at the same elevation). Analyze the complexity of your solutions showing a worst-case input. *(20')*

►**Exercise 191.** An undirected graph G is *bipartite* when its vertices can be partitioned into two sets V_A, V_B such that each edge in G connects a vertex in V_A with a vertex in V_B. In other words,

no two vertices in V_A are adjacent, and no two vertices in V_B are adjacent. To exemplify, see the graphs below.



bipartite | *bipartite* (same as G_1) | *not bipartite* | *bipartite*

Write an algorithm IS-BIPARTITE(G) that takes an undirected graph G and outputs TRUE if and only if G is bipartite. (**Hint:** you may use a simple BFS in which you keep track of which vertex is in which partition.) *(20')*

►**Exercise 192.** Algorithm IS-GOOD(x) classifies a number x as "good" or "not good" in constant time $O(1)$.

Question 1: Write an algorithm GOOD-ARE-ADJACENT(A) that takes a sequence of numbers and, using algorithm IS-GOOD, returns TRUE if all the "good" numbers in A are adjacent, or FALSE otherwise. GOOD-ARE-ADJACENT(A) must not change the input sequence A in any way, may allocate only a constant amount of memory, and must run in time $O(n)$. *(10')*

Question 2: Write an algorithm MAKE-GOOD-ADJACENT(A) that takes a sequence of numbers A and changes A in-place so that all "good" numbers are adjacent. MAKE-GOOD-ADJACENT may allocate only a constant amount of memory and must run in time $O(n)$. *(10')*

►**Exercise 193.** Consider the following decision problem: given a sequence of numbers A and an integer k, returns TRUE if A contains at least k identical values, or FALSE otherwise. Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. Is the problem in P? Write an algorithm that proves it is, or argue the opposite. *(20')*

►**Exercise 194.** Write an algorithm MAXIMAL-COMMON-SUBSTRING(X, Y) that takes two strings X and Y, and returns the maximal length of a common substring of X and Y. For example, if X = 'BDDBADCDCCDCBAD' and Y = 'DDCBCDAABAAC', the output should be 3, since there is a 3-character common substring ('DCB') but no 4-character common substring. Analyze the complexity of your solution. *(20')*

►**Exercise 195.** We say that a node in a binary tree is *unbalanced* when the number of nodes in its left subtree is more than twice the number of nodes in its right subtree plus one, or vice-versa. Write an algorithm BST-COUNT-UNBALANCED-NODES(t) that takes a binary search tree t (the root), and returns the number of unbalanced nodes in the tree. Analyze the complexity of BST-COUNT-UNBALANCED-NODES(t). (**Hint:** an algorithm can return multiple values. For example, the statement **return** x, y returns a pair of values, and if F() returns a pair of values, you can read them with a, b = F().) *(20')*

►**Exercise 196.** Consider the following algorithm that takes an array A of numbers:

```
ALGO-X(A)
1  x = 0
2  for i = 1 to A.length − 1
3      for j = i + 1 to A.length
4          if ALGO-Y(A, i, j) and A[j] − A[i] > x
5              x = A[j] − A[i]
6  return x
```

```
ALGO-Y(A, i, j)
1  for k = i to j − 1
2      if A[k] > A[k + 1]
3          return FALSE
4  return TRUE
```

Question 1: Briefly explain what ALGO-X does and analyze the complexity of ALGO-X by describing a worst-case input. *(10')*

Question 2: Write an algorithm LINEAR-ALGO-X(A) that is equivalent to ALGO-X but runs in linear time. *(20')*

►**Exercise 197.** Let P be an array of points on a plane, each with its Cartesian coordinates $P[i].x$ and $P[i].y$.

Question 1: Write an algorithm FIND-SQUARE(P) that returns TRUE if and only if there are four points in P that form a square. Briefly analyze the complexity of your solution. *(10')*

Question 2: Write an algorithm FIND-SQUARE(P) that solves the problem of Exercise 1 in time $O(n^2 \log n)$. If your solution for Exercise 1 already does that, then simply say so. *(20')*

►**Exercise 198.** Implement a priority queue based on a heap. You must implement the following algorithms:

- INITIALIZE(Q) creates an empty queue. The complexity of INITIALIZE must be $O(1)$.
- ENQUEUE(Q, obj, p) adds an object *obj* with priority p to a queue Q. The complexity of ENQUEUE must be $O(\log n)$.
- DEQUEUE(Q) extracts and returns an object from a queue Q. The returned object must be among the objects in the queue that were inserted with the lowest priority. The complexity of DEQUEUE must be $O(\log n)$.

(**Hint:** Consider Q as an object to which you can add attributes. For example, you may write $Q.A$ = new array, and then later write $Q.A[i]$.) *(20')*

►**Exercise 199.** Implement an algorithm MAXIMAL-DISTANCE(A) that takes an array A of numbers and returns the maximal distance between any two distinct elements in A, or 0 if A contains less than two elements. MAXIMAL-DISTANCE(A) must run in time $O(n)$. *(10')*

►**Exercise 200.** The *height* of a binary tree is the maximal number of nodes on a branch from the root to a leaf node. In other words, it is the maximal number of nodes traversed by a simple path starting at the root. Implement an algorithm BST-HEIGHT(t) that returns the height of a binary search tree rooted at node t. BST-HEIGHT(t) must run in time $O(n)$. *(10')*

►**Exercise 201.** Consider the following decision problem: given a graph $G = (V, E)$ where the edges are weighted by a weight function $w : E \to \mathbb{R}$, and given a number t, output *true* if there is a set of non-adjacent edges $S = \{e_1, e_2, \ldots, e_k\}$ of total weight greater or equal to t, so $\sum w(e_i) \geq t$; or output *false* otherwise. For example, the vertices could represent people, say the students in the Algorithms class, and an edge $e = (u, v)$ with weight $w(e)$ could represent the affinity of the couple (u, v). The question is then, given an affinity value t, tell whether the students in the Algorithms class can form monogamous couples of total affinity value at least t. Argue whether this decision problem is in NP or not, and if it is, then write an algorithm that proves it. *(20')*

►**Exercise 202.** Consider the following game: you are given a set of n valuable objects placed on a 2D plane with non-negative x, y coordinates. In practice, you are given three arrays X, Y, V, such that $X[i]$, $Y[i]$, and $V[i]$ are the x and y coordinates and the *value* of object i, respectively. You start from position 0, 0, and can only move horizontally to the right (increasing your x coordinate) or vertically upward (increasing your y coordinate). Your goal is to reach and collect valuable objects. Write an algorithm MAXIMAL-GAME-VALUE(X, Y, V) that returns the maximal total value you can achieve in a given game. *(30')*

►**Exercise 203.** Write an algorithm MAXIMAL-SUBSTRING(S) that takes an array S of strings, and returns a string x of maximal length such that x is a substring of every string $S[i]$. Also, analyze the complexity of MAXIMAL-SUBSTRING as a function of the size $n = |S|$ of the input array, and the maximal size m of any string in S. *(20')*

►**Exercise 204.** Consider the following algorithm that takes an array A of numbers:

```
ALGO-X(A)
 1  x = 0
 2  y = 0
 3  for i = 1 to A.length
 4       k = 1
 5       for j = i + 1 to A.length
 6            if A[i] == A[j]
 7                 k = k + 1
 8       if x < k
 9            x = k
10            y = A[i]
11  return y
```

Question 1: Briefly explain what ALGO-X does and analyze the complexity of ALGO-X by describing a worst-case input. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X that does the same as ALGO-X but with a strictly better time complexity. Also analyze the complexity of BETTER-ALGO-X. *(10')*

►**Exercise 205.** Write an algorithm GRAPH-DEGREE(G) that takes an undirected graph represented by its adjacency matrix G and computes the *degree* of G. The degree of a graph is the maximal degree of any vertex of G. The degree of a vertex v is the number of edges that are adjacent to v. Also analyze the complexity of GRAPH-DEGREE(G). *(15')*

►**Exercise 206.** Write an algorithm FIND-3-CYCLE(G) that takes an undirected graph represented as an adjacency list, and returns TRUE if G contains a cycle of length 3, or FALSE otherwise. Also, analyze the complexity of FIND-3-CYCLE(G). *(15')*

►**Exercise 207.** Write an algorithm LONGEST-COMMON-PREFIX(S) that takes an array of strings S, and returns the maximal length of a string that is a prefix of at least two strings in S. Also, analyze the complexity of your solution as a function of the size n of the input array S, and the maximal size m of any string in S. For example, with S = ["ciao", "lugano", "bella"] the result is 0, because the only common prefix is the empty string, while with S = ["professor", "prefers", "to", "teach","programming"] the result is 3 because "pro" is a prefix of at least two strings. *(20')*

►**Exercise 208.** Write an algorithm LONGEST-K-COMMON-PREFIX(S, k) that takes an array of strings S and an integer k, and returns the maximal length of a string that is a prefix of at least k strings in S. Also, analyze the complexity of your solution as a function of k, the size n of the input array S, and the maximal size m of any string in S. For example, with S = ["algorithms", "and", "data", "structures"] and $k = 3$, the result is 0, because the only common prefix common to at least three strings is the empty string. While with S = ["professor", "prefers", "to", "teach","programming"] and $k = 3$, the result is 2 because the longest prefix common to at least three strings is "pr". *(20')*

►**Exercise 209.** Consider the following decision problem: given a directed and weighted graph G (with weighted arcs), output TRUE if and only if G contains a path of length 3 and of negative total weight; otherwise output FALSE. Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. Is the problem in P? Write an algorithm that proves it is, or argue the opposite. *(20')*

►**Exercise 210.** Given a collection A of numbers and a number x, the *upper bound* of x in A is the minimal value $a \in A$ such that $x \leq a$, or NULL if no such value exists. For example, given $A = [7, 20, 1, 3, 4, 3, 31, 50, 9, 11]$, the upper bound of $x = 15$ is 20, while the upper bound of $x = 9$ is 9 and the upper bound of $x = 51$ is NULL.

Question 1: Write an algorithm UPPER-BOUND(A, x) that returns the upper bound of x in an array A. Also analyze the complexity of UPPER-BOUND. *(20')*

Question 2: Write an algorithm UPPER-BOUND-SORTED(A, x) that returns the upper bound of x in a sorted array A in time $o(n)$. Analyze the complexity of UPPER-BOUND-SORTED. *(20')*

Question 3: Write an algorithm UPPER-BOUND-BST(T, x) that returns the upper bound of x in a binary search tree T. Analyze the complexity of UPPER-BOUND-BST. *(20')*

►**Exercise 211.** Write an algorithm SUM-OF-THREE(A, s) that takes an array A of n numbers and a number s, and in $O(n^2)$ time decides whether A contains three distinct elements that add up to s. That is, SUM-OF-THREE(A, s) returns TRUE if there are three indexes $1 \le i < j < k \le n$ such that $A[i] + A[j] + A[k] = s$, or FALSE otherwise. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(20')*

►**Exercise 212.** The following algorithm takes an array A of numbers, and a number x:

```
ALGO-X(A, x)
 1  i = A.length
 2  j = 1
 3  while i > 0
 4        if j == i
 5              j = 1
 6              i = i − 1
 7        elseif A[i] − A[j] > x or A[j] − A[i] > x
 8              return TRUE
 9        else j = j + 1
10  return FALSE
```

Question 1: Briefly explain what ALGO-X does and analyze the complexity of ALGO-X by describing a worst-case input. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X(A, x) that is functionally equivalent to ALGO-X but with a strictly better time complexity. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(10')*

►**Exercise 213.** Consider the following algorithm that takes an array A of numbers, and an integer k:

```
ALGO-S(A, k)
1  for i = 1 to A.length
2        if ALGO-R(A, A[i]) == k
3              return A[i]
4  return NULL
```

```
ALGO-R(A, y)
1  c = 0
2  for i = 1 to A.length
3        if A[i] < y
4              c = c + 1
5  return c
```

Question 1: Briefly explain what ALGO-S does and analyze the complexity of ALGO-S by describing a worst-case input. *(10')*

Question 2: Write an algorithm BETTER-ALGO-S(A, k) that is functionally equivalent to ALGO-S(A, k) but with a strictly better complexity. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(10')*

►**Exercise 214.** An array A of n numbers contains only four values, possibly repeated many times. Write an algorithm SORT-SPECIAL(A) that sorts A in-place and in time $O(n)$. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(20')*

►**Exercise 215.** Write an algorithm HEAP-PROPERTIES(A) that takes an array A of n numbers and in $O(n)$ time returns one of four values: -1, if A satisfies the min-heap property; 1, if A satisfies the max-heap property; 2, if A satisfies both the max-heap and min-heap properties; 0, if A does not satisfy either the max-heap or min-heap properties. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(20')*

►**Exercise 216.** You are given a constant-time decision algorithm COMPATIBLE(x, y) that, given two objects x and y tells whether x and y are compatible. The relation expressed by the COMPATIBLE(x, y) algorithm is *symmetric*, meaning that COMPATIBLE(x, y) implies COMPATIBLE(y, x), and *transitive*, meaning that COMPATIBLE(x, y) and COMPATIBLE(y, z) implies COMPATIBLE(x, z). In other words, it is an *equivalence* relation.

Write an algorithm MAX-COMPATIBLE-PAIRING(A) that takes an array of n objects, and in $O(n^2)$ time, returns the maximum number of compatible pairs that can be formed from the objects in A. A compatible pair is a pair of distinct compatible elements, that is, a pair of indexes $1 \le i < j \le n$ such that COMPATIBLE($A[i], A[j]$) == TRUE. Each element (index) may appear in only one pair. Analyze the complexity of your solution and briefly explain the algorithm by commenting on its non-obvious parts. *(20')*

►**Exercise 217.** Consider an infinite chessboard in which the rows and columns are numbered with corresponding integers in their natural order $(\ldots -3, -2, -1, 0, 1, 2, 3, \ldots)$. You are given two arrays W and B of positions of white and black queens, respectively, such that $W[i].row$ and $W[i].col$ are the row and column of the i-th white queen, and correspondingly $B[i].row$ and $B[i].col$ are the row and column of the i-th black queen.
Write an algorithm WHITE-ATTACKS-BLACK(W, B) that takes the two arrays of white and black queens, and returns TRUE if and only if there is a white queen that attacks a black queen. The complexity of your solution must be $o(n^2)$, meaning strictly less than quadratic. (Recall that a queen in row i and column j attacks all positions in row i, all positions in column j, and all positions in the two 45-degree diagonals that pass through the square in row i and column j.) *(20')*

►**Exercise 218.** We say that a node in a binary search tree is *full* if it has both a left and a right child.

Question 1: Write an algorithm called COUNT-FULL-NODES(t) that takes a binary search tree rooted at node t, and returns the number of full nodes in the tree. Analyze the complexity of your solution. *(10')*

Question 2: Write an algorithm called NO-FULL-NODES(t) that takes a binary search tree rooted at node t, and changes the tree in-place, using only rotations, so that the tree does not contain any full node. Analyze the complexity of your solution. *(20')*

►**Exercise 219.** Consider the following decision problem: given two arrays A and B, both containing n numbers, output TRUE if and only if there is a number k and a permutation A' of A such that $A'[i] + B[i] = k$ for all positions $i \in \{1, \ldots, n\}$.

Question 1: Is the problem in NP? Write an algorithm that proves it is, or argue otherwise. *(10')*

Question 2: Is the problem in P? Write an algorithm that proves it is, or argue otherwise. *(20')*

►**Exercise 220.** Write an algorithm called MINIMAL-CONTIGUOUS-SUM(A) that takes an array A of numbers, and outputs the value of the minimal contiguous sub-sequence sum in time $O(n)$. A contiguous sub-sequence sum is the sum of some contiguous elements of A. For example, if A is the sequence

$$-1, 2, -2, -4, 1, -2, 5 - 2 - 3, 1, 2, -1$$

then the minimal contiguous sub-sequence sum is -7, which is the sum of the elements $-2, -4, 1, -2$. *(20')*

►**Exercise 221.** Write an algorithm called HAS-CYCLE(G) that takes a directed graph G represented as an adjacency list, and returns TRUE whenever G contains one or more cycles. You can denote the adjacency list of a vertex v in G as $G.Adj[v]$. Your solution must have a polynomial and possibly linear complexity. Briefly analyze the complexity of your solution. *(20')*

►**Exercise 222.** A DNA sequence S is an array of characters (a string) where each character $S[i]$ is one of 'A', 'C', 'G', or 'T'. Write an algorithm DNA-PERMUTATION-SUBSTRING(S, X) that takes a large DNA sequence S and a smaller sequence X, and in linear time returns TRUE if and only if S contains a contiguous subsequence (a substring) that is a permutation of X. For example, DNA-PERMUTATION-SUBSTRING("GCCATCAGTGACGAAGCT", "TAGG") would return TRUE, since the long sequence contains the contiguous subsequence "AGTG", which is a permutation of the sequence "TAGG". *(30')*

►**Exercise 223.** Consider the following algorithm that takes a non-empty array A of numbers:

```
ALGO-X(A)
 1  n = A.length
 2  let B be an array of size n
 3  for i = 1 to n
 4      B[i] = 0
 5  m = 1
 6  x = A[1]
 7  for i = 1 to n
 8      if B[i] == 0
 9          B[i] = 1
10          for j = i + 1 to n
11              if A[i] == A[j]
12                  B[i] = B[i] + 1
13                  B[j] = 1
14          if m < B[i] or (m == B[i] and x > A[i])
15              x = A[i]
16              m = B[i]
17  return x
```

Question 1: Briefly describe what ALGO-X does and analyze the complexity of ALGO-X. *(10')*

Question 2: Write an algorithm called BETTER-ALGO-X that does exactly the same thing, but with a strictly better asymptotic complexity. Analyze the complexity of BETTER-ALGO-X. *(20')*

►**Exercise 224.** Consider the problem of comparing two binary search trees.

Question 1: Write an algorithm BST-EQUALS(t_1, t_2) that takes the roots t_1 and t_2 of two binary search trees and returns TRUE if and only if the tree rooted t_1 is exactly the same as the tree rooted at t_2, meaning that the two trees have nodes with the same keys connected in exactly the same way. Also, analyze the complexity of your solution. *(10')*

Question 2: Write an algorithm BST-EQUAL-KEYS(t_1, t_2) that takes the roots t_1 and t_2 of two binary search trees and returns TRUE if and only if the tree rooted t_1 contains exactly the same keys as the tree rooted at t_2. *(20')*

►**Exercise 225.** Consider an infinite chessboard in which the rows and columns are numbered with corresponding integers in their natural order $(\ldots -3, -2, -1, 0, 1, 2, 3, \ldots)$. Write an algorithm KNIGHT-DISTANCE(r_1, c_1, r_2, c_2) that takes two positions on the chessboard, identified by the respective row and column numbers, and returns the minimal number of hops it would take a knight to go from the first position to the second position. Also, analyze the complexity of your solution. *Hints:* a knight moves in a single hop by two squares horizontally and by one square vertically, or vice-versa. Notice that what matters is the *distance*, not the absolute positions, so consider computing the distance between any position (r, c) and the $(0, 0)$ position. Consider a dynamic-programming solution. Also notice that the problem has symmetries that can greatly simplify the solution. For example, the distance from $(0, 0)$ to position (a, b) is the same as to position (b, a). *(30')*

►**Exercise 226.** Consider a directed graph G of 20 vertexes, numbered from 1 to 20, and defined by the following adjacency list

$v \to adj(v)$
$1 \to 2$
$2 \to 8\ 9$
$3 \to 2\ 4\ 5\ 6$
$4 \to 10\ 11\ 12\ 13\ 14\ 15\ 5\ 9$
$5 \to 18\ 7$
$6 \to 5\ 7$
$7 \to 18\ 19\ 4$
$8 \to 9$
$9 \to 10$
$10 \to 11$
$11 \to 12\ 14$
$12 \to 14$
$13 \to 14\ 17\ 20$
$15 \to 13\ 16\ 5$
$16 \to 13\ 17\ 5$
$17 \to 18\ 19$
$18 \to 19$
$20 \to 14\ 17$

(**Hint:** draw the graph and use the drawing to answer the following questions.)

Question 1: Compute a depth-first search on G. Write the three vectors P, D, and F that, for each vertex, hold the *previous vertex* in the depth-first forest, the *discovery time*, and the *finish time*, respectively. Whenever necessary, iterate through vertexes in numeric order. *(20')*

Question 2: Compute a breadth-first search on G starting from vertex 1. Write the two vectors P and D that, for each vertex, hold the *previous vertex* in the breadth-first tree and the *distance*, respectively. Whenever necessary, iterate through vertexes in numeric order. *(20')*

►**Exercise 227.** Consider the following decision problem: given a sequence A of numbers and given an integer k, return TRUE if and only if A contains either an increasing or a decreasing subsequence of length k. The elements of the subsequence must maintain their order in A but do not have to be contiguous. For example, $A = [4, 5, 3, 8, 3, 9]$ contains an increasing sequence of length $k = 4$ (4, 5, 8, 9), but neither an increasing or decreasing sequence of length $k = 5$.

Question 1: Is the problem in NP? Write an algorithm that proves it is, or argue the opposite. *(10')*

Question 2: Is the problem in P? Write an algorithm that proves it is, or argue the opposite. *(20')*

►**Exercise 228.** Given a sequence of numbers $A = \langle a_1, a_2, \ldots, a_n \rangle$, we define a *maximal contiguous subsequence* as a contiguous subsequence of numbers in A, starting at position i and ending at position j with $1 \leq i \leq j \leq n$, whose sum is maximal.

Question 1: Write an algorithm MCS-VALUE(A) that, given a sequence A, returns the sum of a maximal contiguous subsequence in A. Also, analyze the complexity of your solution. *(10')*

Question 2: Write an algorithm MCS-VALUE-LINEAR(A) that, given a sequence A, returns the sum of a maximal contiguous subsequence in A with $O(n)$ complexity. *(20')*

►**Exercise 229.** Analyze the following algorithms that take an array A of integers. First, briefly describe what the algorithm does, and then analyze the best- and worst-case complexity as functions of the size of the input $n = |A|$. Your characterizations should be as tight as possible. Briefly justify your answers.

Question 1: Describe and analyze the following ALGO-X *(10')*

```
ALGO-X(A)
1  for i = |A| downto 2
2      s = TRUE
3      for j = 2 to i
4          if A[j − 1] > A[j]
5              swap A[j − 1] ↔ A[j]
6              s = FALSE
7      if s == TRUE
8          return
```

Question 2: Describe and analyze the following ALGO-Y *(10')*

```
ALGO-Y(A)
1  i = 1
2  j = |A|
3  while i < j
4      if A[i] > A[j]
5          swap A[i] ↔ A[i + 1]
6          if i + 1 < j
7              swap A[i] ↔ A[j]
8          i = i + 1
9      else j = j − 1
```

►**Exercise 230.** Write an algorithm PARTITION-ZERO(A) that takes an array of numbers A and, in $O(n)$ time, rearranges the elements of A in-place so that all the negative elements of A precede all the elements equal to zero that precede all the positive elements. For example, with an initial array $A = [2, 5, 0, -1, 3, -7, 0, 3, -1, 10]$, a valid (but not unique) result of PARTITION-ZERO(A) would be the permuted array $A = [-1, -7, -1, 0, 0, 2, 5, 3, 3, 10]$. *(20')*

►**Exercise 231.** Implement a priority queue. Given two objects x and y, you can test whether x has a higher priority than y by testing the condition $x > y$. Briefly describe the data structure (data and meta-data) and then write three algorithms: PQ-INIT(n) creates, initializes, and returns a priority queue Q of maximal size n; PQ-ENQUEUE(Q, x) enqueues an object x into queue Q; PQ-DEQUEUE(Q) extracts and returns an object x such that there is no other object y in Q such that $y > x$. Both PQ-ENQUEUE and PQ-DEQUEUE must have a complexity $O(\log n)$. *(30')*

►**Exercise 232.** Consider the following algorithm ALGO-X(A, B) that takes two arrays of numbers

```
ALGO-X(A, B)
 1  C = copy of array B
 2  n = C.length
 3  for i = 1 to A.length
 4      j = 1
 5      while j ≤ n
 6          if A[i] == C[j]
 7              swap C[j] ↔ C[n]
 8              n = n − 1
 9          else j = j + 1
10  if n == 0
11      return TRUE
12  else return FALSE
```

Question 1: Briefly explain what ALGO-X does, and analyze its complexity by also describing a worst-case input. *(10')*

Question 2: Write a an algorithm BETTER-ALGO-X that is functionally identical to ALGO-X but with a strictly better time complexity. *(20')*

►**Exercise 233.** Consider the following algorithm QUESTIONABLE-SORT(A) that takes an array of numbers A and intends to sort it in-place.

```
QUESTIONABLE-SORT(A)
1  for i = 1 to A.length − 1
2      for j = i + 1 to A.length
3          if A[i] > A[j]
4              swap A[i] ↔ A[j]
```

Question 1: Is QUESTIONABLE-SORT correct? If so, explain how the algorithm works. If not, show a counter-example. *(10')*

Question 2: Write an algorithm BETTER-SORT that sorts in-place with a strictly better average-case complexity than QUESTIONABLE-SORT. *(10')*

►**Exercise 234.** Write an algorithm LOWER-BOUND(A, x) that takes a sorted array A of numbers and, in $O(\log n)$ time returns the least (smallest) number a_i in A such that $a_i \geq x$. If no such value exists, LOWER-BOUND(A, x) must return a "not-found" error. *(20')*

►**Exercise 235.** Write an algorithm CONTAINS-SQUARE(A) that takes an $\ell \times \ell$ matrix A of numbers, and returns TRUE if and only if A contains a square pattern of equal numbers, that is, a set of equal elements $A_{x,y}$ whose positions, interpreted as points with Cartesian coordinates (x, y), lay on the perimeter of a square. A square pattern consists of at least four elements, so a single number is not a valid square pattern. For example, the following matrix contains a square pattern consisting of elements with value 3. Notice in fact that there are two such square patterns.

$$\begin{bmatrix} 7 & 8 & 3 & 8 & 8 & 3 \\ 7 & 8 & 3 & 3 & 3 & 3 \\ 1 & 3 & 3 & 5 & 8 & 3 \\ 7 & 6 & 3 & 5 & 3 & 3 \\ 0 & 4 & 3 & 3 & 3 & 3 \\ 9 & 9 & 1 & 3 & 7 & 3 \end{bmatrix}$$

Also, analyze the complexity of your solution as a function of $n = \ell^2$. *(20')*

►**Exercise 236.** Write an algorithm MIN-HEAP-CHANGE(H, i, x) that takes a min-heap H of size n, an index i, and a value x, and changes the value $H[i]$ to x, possibly adjusting the heap so as to maintain the min-heap property. MIN-HEAP-CHANGE must run in $O(\log n)$ time. Analyze the complexity of your solution. *(20')*

►**Exercise 237.** Write an algorithm BST-SUBSET(T_1, T_2) that takes two binary search trees T_1 and T_2 (the roots) and returns TRUE if and only if T_1 contains a subset of the keys in T_2. Your solution must run in time $O(n)$, where n is the total size of the two input trees. Analyze the complexity of your solution. *(20')*

►**Exercise 238.** Consider the following decision problem: given a graph $G = (V, E)$ and an integer k, return TRUE if G contains a cycle of length k, or otherwise FALSE. Is this problem in NP? Show a proof of your answer. *(20')*

►**Exercise 239.** Consider the following decision problem: given a graph $G = (V, E)$, return TRUE if G contains a cycle of length 4, or otherwise FALSE. Is this problem in P? Show a proof of your answer. *(20')*

►**Exercise 240.** Write an algorithm SUMS-ONE-TWO-THREE(n) that takes an integer n and, in time $O(n)$, returns the number of possible ways to write n as a sum of 1, 2, and 3. For example, SUMS-ONE-TWO-THREE(4) must return 7 because there are 7 ways to write 4 as a sum of ones, twos, and threes ($1 + 1 + 1 + 1$, $1 + 1 + 2$, $1 + 2 + 1$, $2 + 1 + 1$, $2 + 2$, $1 + 3$, $3 + 1$). Analyze the complexity of your solution. *Hint:* use dynamic programming. *(20')*

►**Exercise 241.** Write an algorithm TWO-PRIMES(n) that takes a number n and returns TRUE if and only if n is the sum of two primes. For example, TWO-PRIMES(12) returns TRUE because $12 = 5 + 7$,

and 5 and 7 are primes, but TWO-PRIMES(11) returns FALSE, because it can not be expressed as the sum of two primes. Analyze the complexity of your solution as a function of n.
Recall that a prime p is a positive integer that can not be written as the product of two positive integers smaller than p. Thus $2, 3, 5, 7, 11, \ldots$ are primes, but 1 and 4 are not. *(20')*

►**Exercise 242.** Consider the following algorithm ALGO-X(A) that takes a non-empty array A of objects each with two numeric attributes: *weight* and *category*.

```
ALGO-X(A)
 1  c = A[1].category
 2  w = −∞
 3  for i = 1 to A.length
 4       t = 0
 5       for j = 1 to A.length
 6            if A[j].category == A[i].category
 7                 t = t + A[j].weight
 8       if t > w or (t == w and c > A[i].category)
 9            c = A[i].category
10            w = t
11  return c
```

Question 1: Describe at a high-level what ALGO-X does, and analyze its complexity. *(10')*

Question 2: Write an algorithm BETTER-ALGO-X that is functionally equivalent to ALGO-X but with a strictly better time complexity. *(20')*

►**Exercise 243.** Consider a min-heap represented internally as an array H with an additional attribute $H.heap\text{-}size$ representing the number of elements in the heap.

Question 1: Write an algorithm MIN-HEAP-INSERT(H, x) that takes a valid min-heap H and inserts a new value x in H. Analyze the complexity of your solution. *(10')*

Question 2: Write an algorithm MIN-HEAP-DEPTH(H) that computes the depth of a given min-heap in $O(\log n)$ time. *(10')*

►**Exercise 244.** Consider the following algorithm ALGO-Y(A) that takes an array A of numbers.

```
ALGO-Y(A)
1  m = −∞
2  for i = 1 to A.length − 1
3       for j = i + 1 to A.length
4            if A[i] + A[j] > m
5                 m = A[i] + A[j]
6  return m
```

Question 1: Describe at a high-level what ALGO-Y does and analyze its complexity. *(10')*

Question 2: Write an algorithm BETTER-ALGO-Y that is functionally equivalent to ALGO-Y and that runs in $O(n)$ time. *(20')*

►**Exercise 245.** Consider the following decision problem: given an array A of numbers, a number m, and an integer k, output TRUE if and only if A contains k distinct elements $A[i_1], A[i_2], \ldots, A[i_k]$ such that $A[i_1] + A[i_2] + \cdots + A[i_k] \geq m$. Is this problem in P? Show a proof of your answer. *(20')*

Solutions

WARNING: solutions are very sparse, meaning that many are missing, most of the solutions are only sketched at a high level, and many may be incorrect! Please, consider contributing your solutions, including alternative solutions, and please report any error you might find to the author (Antonio Carzaniga <antonio.carzaniga@usi.ch>).

▷ *Solution 20.6*
Yes, the exact-change problem is in NP. There is in fact a verification algorithm that, given an instance of the problem (V, x) and a "witness" set S that shows that the solution is 1, can check in polynomial time that S indeed proves that the solution is 1. Below is such an algorithm:

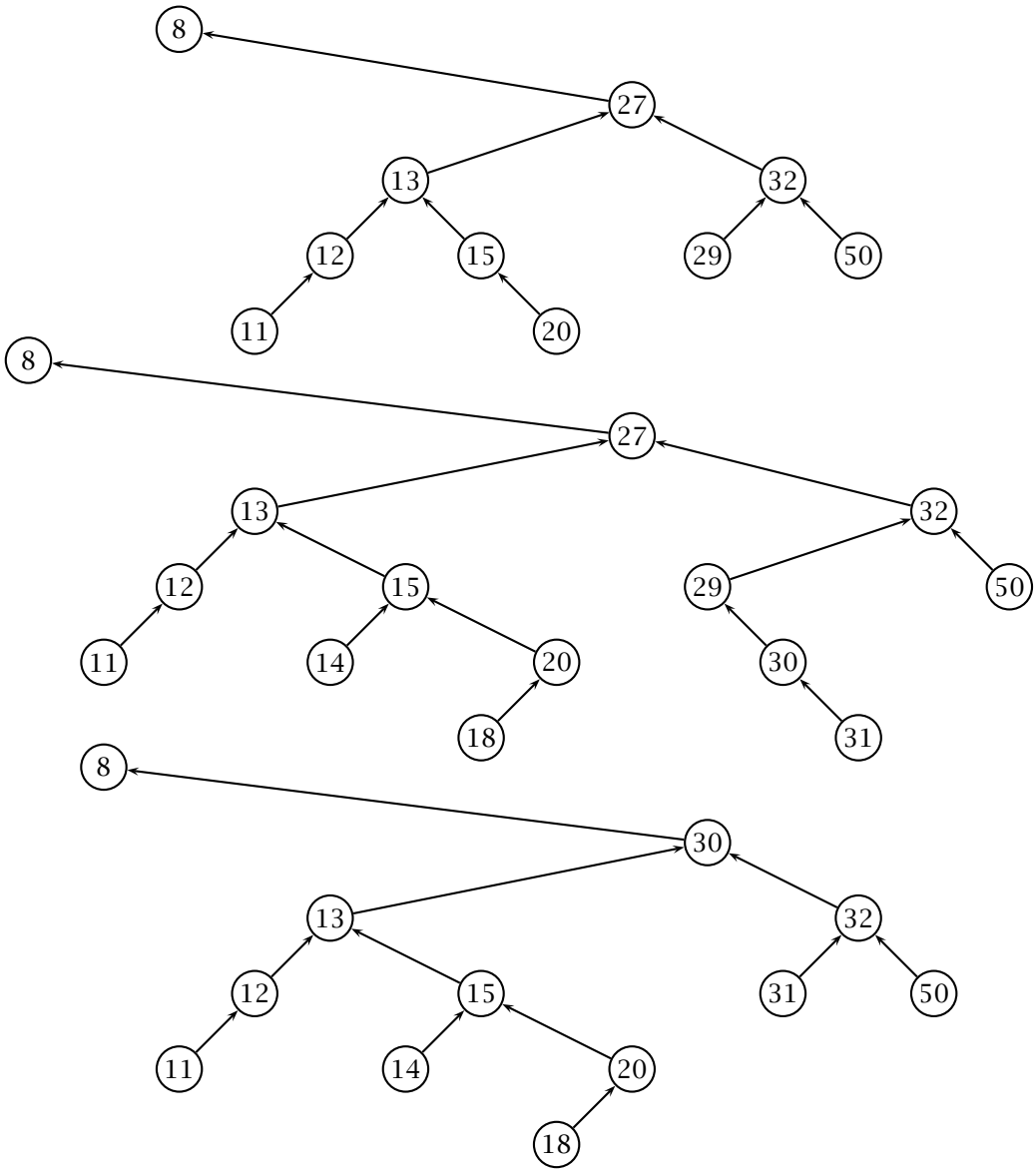
EXACT-CHANGE-VERIFY(V, x, S)

```
1  t = 0
2  for v ∈ S
3        if v ∉ V
4              return FALSE
5        t = t + v
6  if t == x
7        return TRUE
8  else return FALSE
```

▷ *Solution 54*
Quick-sort. Best-case is $O(n \log n)$, worst-case is $O(n^2)$.

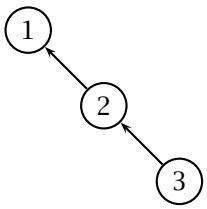
▷ *Solution 55*
ALGORITHM-I sorts the input array in-place. In the best case, the algorithm terminates in the first execution of the outer loop, with the condition s == TRUE. This is the case when the inner loop does not swap a single element of the array, meaning that the array is already sorted. So, the best-case complexity is $O(n)$. Conversely, the worst case is when each iteration of the outer loop swaps at least one element. This happens when the array is sorted in reverse order. So, the worst-case complexity is $O(n^2)$.
ALGORITHM-II sorts the input array in-place so that the value $v = A[0]$, that is the element originally at position 0, ends up in position q, and every other element less than v ends up somewhere in $A[1 \ldots q-1]$, that is to the left of q, and every other element less than or equal to v ends up somewhere in $A[q+1 \ldots |A|]$. In other words, ALGORITHM-II partitions the input array in-place using the first element as the "pivot". The loop closes the gap between i and j, which are initially the first and last position in the array, respectively. Each iteration either moves i to the right or j to the left, so each iteration reduces the gap by one. Therefore, in any case—worst case is the same as the best case—the complexity is $O(n)$.

▷ *Solution 67*



▷ *Solution 68*
a) 50 32 20 29 15 13 12 8 27 11
b) 51 43 50 29 32 20 12 8 27 11 15 13
c) 32 29 20 27 15 13 12 8 11

▷ *Solution 69*
Proof: Let $H = [1, 2, 3]$, then T would look like this:



▷ *Solution 73.1*
True. $O(n!)$ is at most $Kn!$ so $\log(Kn!) = \log K + \log 1 + \log 2 + \cdots + \log n \leq n \log n$

▷ *Solution 73.2*
False. as a counter example, let $f(n) = \sqrt{n}$

▷ *Solution 73.3*
False. Counter-example: $f(n) = 1$ and $g(n) = n$.

▷ *Solution 73.4*
False. Counter example: $f(n) = n$ and $g(n) = n$

▷ *Solution 73.5*
False. Counter example: $f(n) = \sqrt{n}$ and $g(n) = \sqrt{(}n)$

▷ *Solution 74*
SHUFFLE-A-BIT has the same common structure as a best-case run of QUICK-SORT. There is an initial linear phase, and then there are two recursions on arrays of size $n/2$. This results in $\log n$ levels of recursion, each having a total cost of $O(n)$. Therefore the complexity is $n \log n$.

▷ *Solution 75.1*
yes.

▷ *Solution 75.2*
yes.

▷ *Solution 75.3*
undefined.

▷ *Solution 75.4*
yes.

▷ *Solution 75.5*
undefined.

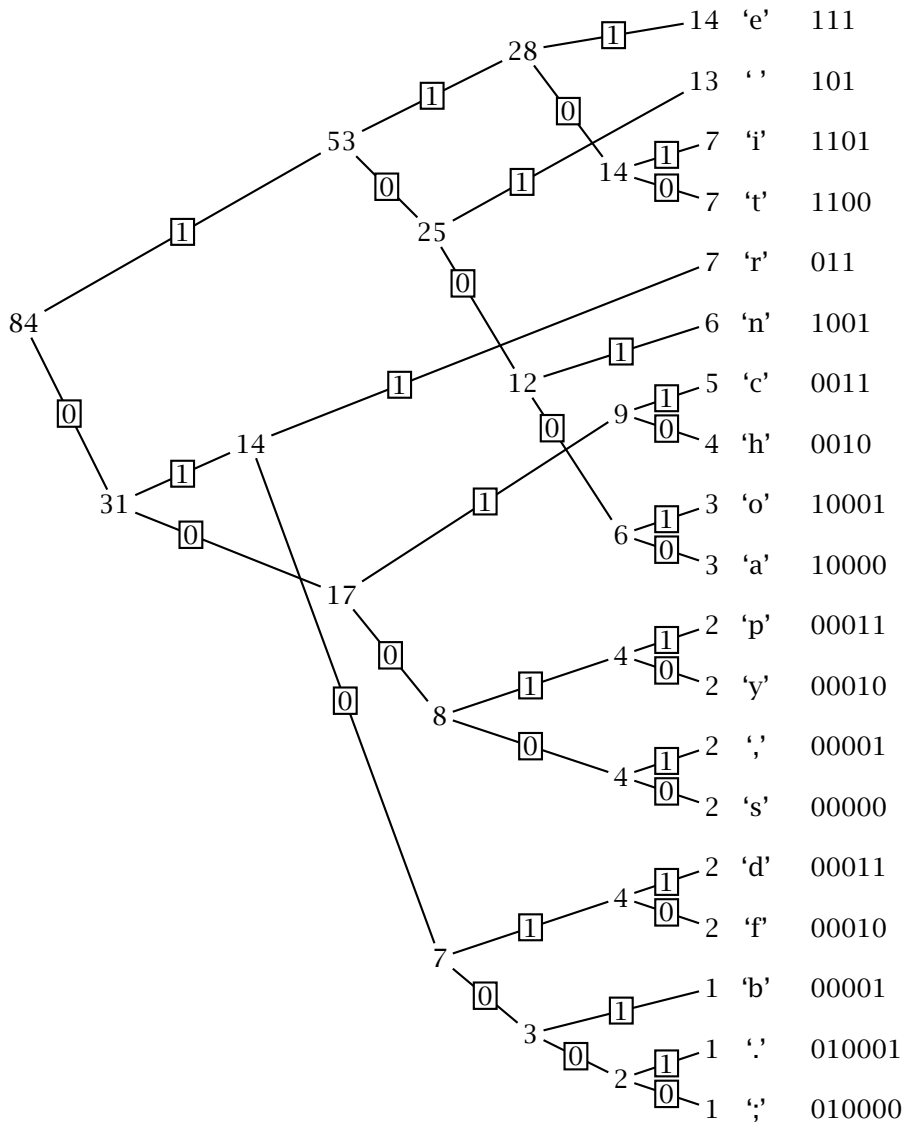
▷ *Solution 75.6*
undefined.

▷ *Solution 75.7*
yes.

▷ *Solution 75.8*
undefined.

▷ *Solution 75.9*
undefined.

▷ *Solution 76*
First figure out the frequencies and sort the characters by frequency. Then we proceed with the derivation:



▷ *Solution 78*

```
ISCOLORVALID(G = (V, E), v)
 1  for each u adjacent to v
 2       if color[u] = color[v]
 3            return FALSE
 4  return TRUE

COLOR(G = (V, E))
 1  for each v ∈ V
 2       color[v] = 0
 3  for each v ∈ V
 4       color[v] = 1
 5       while ISCOLORVALID(G = (V, E), v) = FALSE
 6            color[v] = color[v] + 1
 7  return color
```

▷ *Solution 84*

Given an array A of number, ALGO-X(A) returns TRUE if and only if there are three numbers $x \leq y \leq z \in A$ such that $y - x = z - y$. ALGO-X does that by testing each triple of distinct elements of A. There are $\binom{n}{3} = n(n-1)(n-2)/3!$ such triples, so the complexity is $\Theta(n^3)$.

A better way to do the same thing is as follows:

```
BETTER-ALGO-X(A)
 1  sort A
 2  for i = 1 to A.length − 2
 3      for j = i + 2 to A.length
 4          m = (A[i] + A[j])/2
 5          if BINARY-SEARCH(A[i + 1 . . . j − 1], m)
 6              return TRUE
 7  return FALSE
```

In essence, after sorting the numbers, this algorithm tests each pair of non-adjacent numbers and then looks for the median using a binary search. There are $O(n^2)$ pairs of non-adjacent numbers in A, and binary-search costs $O(\log n)$, so the complexity is $O(n^2 \log n)$.

▷ *Solution 95*

```
TREE-TO-VINE(t)
 1  if t == NIL
 2      return (NIL
 3  while t.left ≠ NIL
 4      t = BST-RIGHT-ROTATE(t)
 5  root = t
 6  while t.right ≠ NIL
 7      while t.right.left ≠ NIL
 8          t.right = BST-RIGHT-ROTATE(t.right)
 9      t = t.right
10  return root
```

The best-case complexity is $\Theta(n)$, which corresponds to the case of BST that is already a vine. The general worst-case complexity is certainly $O(n^2)$, since the outer loop (line 6) can run for at most n iterations, and similarly the inner loop (line 7) can also run for at most n iterations. However, it is not immediately obvious that the quadratic complexity is "tight". In fact, the complexity is $\Theta(n)$ also in the worst case. To see why, consider what happens to an edge e in the tree. If e is a left edge—that is, an edge connecting a parent node to a left child node—then at some point the algorithm will rotate e with a right rotation of the parent node, transforming e into a right edge that the algorithm will then simply traverse once. In other words, a left edge will involve two steps: a right rotation plus a traversal. If e is a right edge, then e might be immediately traversed, or it might be transformed into a left edge due to a rotation of another edge right above and to the right of e, which means that at some point e will be treated like any other left edge, so rotated and then traversed. In any case, every edge induces a constant-time process, which means that the algorithm linear, since there are $n - 1$ edges in the tree.

▷ *Solution 96*

```
IS-PERFECTLY-BALANCED(t)
 1  if t == NIL
 2      return (TRUE, 0)
 3  (balanced_l, weight_l) = IS-PERFECTLY-BALANCED(t.left)
 4  (balanced_r, weight_r) = IS-PERFECTLY-BALANCED(t.right)
 5  if balanced_l and balanced_l and |weight_r − weight_l| ≤ 1
 6      return (TRUE, weight_l + weight_r + 1)
 7  else return (FALSE, weight_l + weight_r + 1)
```

▷ *Solution 98*

We are not allowed to modify H, and we are not allowed to create a copy of H that we can then sort. So, we must print the elements in order, by simply reading H. We know that each number is unique in H, so the idea is this: we start from the minimum value in H, which happens to be in the first position of H, print that value, and then look for the second-smallest number, which we can simply find with a linear scan. We then proceed with the third-smallest, and so on, which again we can find with a linear scan. Notice that we can use a linear scan to find the i-th smallest element by simply considering only those elements in H that are greater than the smallest element we found in the previous $(i - 1)$ scan.

In pseudo-code:

HEAP-PRINT-IN-ORDER(H)

```
1  m = H[0]
2  print m
3  for i = 2 to H.length
4       x = ∞
5       for j = 2 to H.length
6            if H[j] > m and H[j] < x
7                 x = H[j]
8       m = x
9       print m
```

It is easy to see that the complexity of HEAP-PRINT-IN-ORDER is $\Theta(n^2)$.

▷ *Solution 102*
ALGO-X removes every element equal to k from array A. with a complexity of $\Theta(n^2)$.
Consider as a worst-case input an array A in which all n values are equal to k. In this case, ALGO-X would iterate over lines 3 and 4 (always with i equal to 1). In each iteration, ALGO-X would then invoke ALGO-Y (again with i equal to 1), which would then iterate over the length of the array, effectively removing the i-th element by shifting every subsequent element to the left by 1 position, and then by cutting the length of the array by 1.
So, ALGO-Y would run for n iterations the first time, then $n-1$ the second time, then $n-2$, and so on, until the array is completely empty. The complexity is therefore $n+(n-1)+\ldots+2+1=\Theta(n^2)$.
A better way to remove every element equal to k from an array A is as follows.
BETTER-ALGO-X(A, k)

```
1  j = 1
2  for i = 1 to A.length
3       if A[i] ≠ k
4            A[j] = A[i]
5            j = j + 1
6  A.length = j − 1
```

▷ *Solution 106.1*
No.

▷ *Solution 106.2*
No. SAT is NP-complete, meaning that every problem in NP can be reduced to SAT (polynomially), so if SAT can then be reduced to Q, then that measnt that all problems in NP can be reduced to Q, which makes Q an NP-hard problem. However, to say that Q is NP-complete, we also have to know that Q is itself in NP.

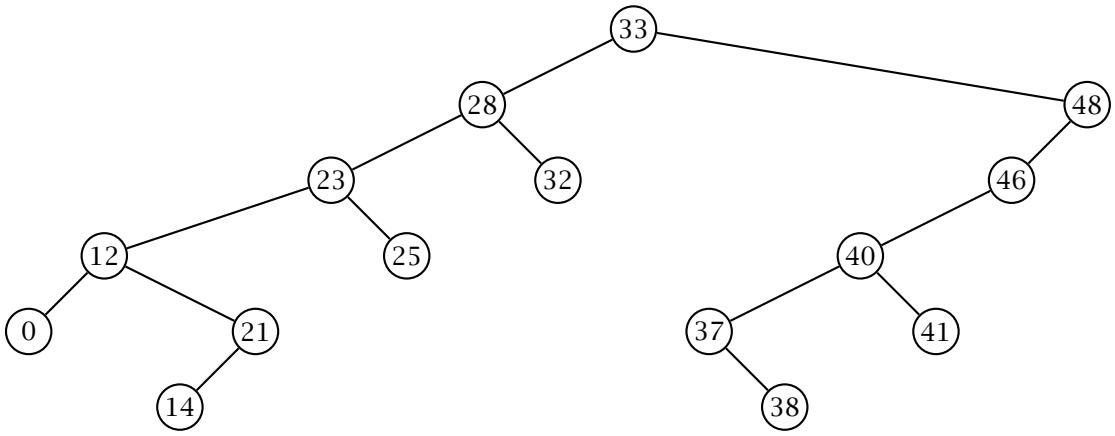
▷ *Solution 106.3*
Yes. SAT is in NP, and therefore there is a polynomial-time *verification* algorithm for SAT. Since Q (and Q') can be transformed into SAT, that means that one can implement a polynomial verification algorithm also for Q (and Q'), by first transforming the Q instance into an instance of SAT, and then running the verification algorithm for SAT. Thus both Q and Q' are polynomially verifiable.

▷ *Solution 106.4*
No. We can say that Q is not more complex than Q', but we can not say much about Q'.

▷ *Solution 106.5*
Yes. Q is polynomially solvable, since it can be transformed into Q' and then solved in polynomial time through algorithm A. This means that Q is in P and therefore also in NP.

▷ *Solution 106.6*
Yes. Q is in P, since it can be easily solved through a breadth-first search.

▷ *Solution 133.2*



Optimal sequence: 32, 21, 25, 40, 37, 46, 41, 12, 23, 48, 14, 33, 38, 0, 28.

▷ *Solution 137*

We can first start by modeling MAXIMAL-NON-ADJACENT-SEQUENCE-WEIGHT as a classic recursive dynamic-programming algorithm. Given a sequence $A = a_1, a_2, \ldots, a_n$, there are two cases:

(i) the maximal sequence includes a_1, and therefore does not include a_2 and instead includes the maximal sequence for the remaining subsequence $a_3 \ldots a_n$;

(ii) the maximal sequence does not include a_1 and therefore is the same as the maximal sequence for the subsequence starting at a_2.

Thus the maximal solution is the best of these two. Let $OPT(a_1, a_2, \ldots, a_n)$ denote the maximal weight of non-adjacent elements from a sequence $a_1, a_2, \ldots, a_n$. With this, the algorithm is as follows:

$$OPT(a_1, a_2, a_3, \ldots, a_n) = \max\{a_1 + OPT(a_3, \ldots, a_n), OPT(a_2, \ldots, a_n)\}$$

Now we just have to write this simple, recursive dynamic-programming solution as a single iteration. This can be done by remebering only two values in each iteration, namely the optimal value for the previous two elements in the sequence. We can perform this iteration in either direction, so here we do it in increasing order, left-to-right. Therefore, for each element a_i, we must remember the two previous optimal values $OPT(a_1, \ldots, a_{i-1})$ and $OPT(a_1, \ldots, a_{i-2})$. The full algorithm is as follows:

MAXIMAL-NON-ADJACENT-SEQUENCE-WEIGHT(A)

```
1  p = 0
2  q = 0
3  r = 0
4  for i = 1 to A.length
5       r = max{A[i] + p, q}
6       p = q
7       q = r
8  return r
```

▷ *Solution 152.1*

min, max, min−1, max−1, min−2, max−2, ...

▷ *Solution 152.2*

Dynamic programming: with i going from left to right, let $x(i)$ be the value of the maximal contiguous sequence *ending at position i*. So, $x(1) = A[1]$, $x(i) = \max\{A[i] + x(i-1), A[i]\}$.

▷ *Solution 160*

MAX-HEAP-INSERT(H, k)

```
1  H.heap-size = H.heap-size + 1
2  H[H.heap-size] = k
3  i = H.heap-size
4  while i > 1 and H[i] > H[⌊i/2⌋]
5       swap H[i] ↔ H[⌊i/2⌋]
6       i = ⌊i/2⌋
```

The complexity is $\Theta(\log n)$.

▷ *Solution 161.1*

FIND-ELEMENTS-AT-DISTANCE(A, k)

```
1  for i = 1 to A.length
2       if BINARY-SEARCH(A[i + 1 ... A.length], k − A[i])
3            return TRUE
4  return FALSE
```

The complexity is $\Theta(n \log n)$, since for each of the n elements, we perform a binary search that runs in $\Theta(\log n)$.

▷ *Solution 161.2*

FIND-ELEMENTS-AT-DISTANCE(A, k)

```
1  i = 1
2  j = 2
3  while j ≤ A.length
4       if A[j] − A[i] < k
5            j = j + 1
6       elseif A[j] − A[i] > k
7            i = i + 1
8       else return TRUE
9  return FALSE
```

In each iteration of the loop we either increase j or i by one (or we return). Also, the loop is such that $j \geq i$, so in at most $\Theta(n)$ iterations we push j beyond $A.length$. Thus the complexity is $\Theta(n)$.

▷ *Solution 162*

IS-PRIME(x)

```
1  i = 2
2  while i * i < x
3       if i divides x
4            return TRUE
5       i = i + 1
6  return FALSE
```

PARTITION-PRIMES-COMPOSITES(A)

```
1  i = 1
2  j = 1
3  while i < j
4       if IS-PRIME(A[j])
5            swap A[j] ↔ A[i]
6            i = i + 1
7       elseif not IS-PRIME(A[i])
8            swap A[j] ↔ A[i]
9            j = j − 1
10      else i = i + 1
11           j = j − 1
```

IS-PRIME runs in $\Theta(\sqrt{m})$, while PARTITION-PRIMES-COMPOSITES requires $\Theta(n)$ basic operations and $\Theta(n)$ invocations of IS-PRIME. The complexity is therefore $\Theta(n\sqrt{m})$.

▷ *Solution 163*

In this exercise, randomization or rotations cannot be used to balance the height of the BST. So, input sequence A must be pre-sorted so that, inserting elements in the tree in the new order, the resulting BST has still minimal height, $O(\log n)$, even using the classic insertion algorithm (that

could potentially result in unbalanced trees). Intuitively, this is possible by inserting elements in this order: $median(1, n)$, $median(1, \frac{n}{2})$, $median(\frac{n}{2}, n)$, $median(1, \frac{n}{4})$, $median(\frac{n}{4}, \frac{n}{2})$, $median(\frac{n}{2}, \frac{3n}{4})$, $median(\frac{3n}{4}, n)$. Or, equivalently, $median(1, n)$, $median(1, \frac{n}{2})$, $median(1, \frac{n}{4})$, $median(\frac{n}{4}, \frac{n}{2})$, $median(\frac{n}{2}, n)$, $median(\frac{n}{2}, \frac{3n}{4})$, $median(\frac{3n}{4}, n)$. The input array can be sorted in this order by using the functions below:

```
SORT-FOR-BALANCED-BST(A)
1  sort A in non-descending order
2  PRINT-R(A, 1, A.length)
```

```
PRINT-R(A, i, j)
1  if i ≤ j
2      m = ⌊(i + j)/2⌋
3      print A[m]
4      PRINT-R(A, i, m − 1)
5      PRINT-R(A, m + 1, j)
```

PRINT-R runs in $O(n)$, since it simply prints one element—the median element, since the input is sorted—and then recurses on the left and side parts by excluding the element it just printed. In the end, PRINT-R runs (recursively) exactly once for each element of the array. So, the complexity of PRINT-R is $O(n)$ and the dominating cost for SORT-FOR-BALANCED-BST is the cost of sorting, which can be done in $O(n \log n)$.

▷ *Solution 164.1*

```
MINIMAL-SIMPLIFIED-SEQUENCE(A)
1  X = ∅
2  sort A in non-decreasing order
3  for i = A.length downto 3
4      for j = A.length downto 3
5          if BINARY-SEARCH(A[1 . . . j − 1], A[i] − A[j]) ≠ TRUE
6              X = X ∪ {A[i]}
7  return X
```

Hey, is the solutions above incorrect? An alternative solution is below:

```
MINIMAL-SIMPLIFIED-SEQUENCE(A)
1  X = ∅
2  sort A in non-decreasing order
3  for i = 1 to A.length − 1
4      for j = i + 1 to A.length
5          i = BINARY-SEARCH(A[j + 1 . . . A.length], A[i] + A[j])
6          if i > 0
7              X = X ∪ {A[i]}
8  return X
```

The complexity is $\Theta(n^2 \log n)$.

▷ *Solution 164.2*

```
MINIMAL-SIMPLIFIED-SEQUENCE(A)
 1  B = array of A.length zeroes
 2  sort A in non-decreasing order
 3  for i = 1 to A.length − 2
 4       j = i + 1
 5       k = i + 2
 6       while k ≤ A.length
 7            if A[k] − A[j] < A[i]
 8                 k = k + 1
 9            elseif A[k] − A[j] > A[i]
10                 j = j + 1
11            else B[k] = 1
12                 k = k + 1
13  X = ∅
14  for i = 1 to A.length
15       if B[i] == 0
16            X = X ∪ {A[i]}
17  return X
```

▷ *Solution 165*

The algorithm consists of two nested loops. The outer loop takes variable a from n to 1 by dividing a in half at every iteration. Therefore, the values of a are $n, n/2, n/4, n/8 \ldots$. That is, at iteration i of the outer loop, $a = n/2^i$. The outer loop terminates when $n/2^i \leq 1$, that is, it runs for $\lceil \log n \rceil$ iterations.

The inner loop takes variable b from 1 to a^2 by doubling b at every iteration. Therefore the values of b are $1, 2, 4, \ldots$, that is, $b = 2^j$ at the j-th iteration of the inner loop. Therefore the inner loop runs for $2 \log a$ iterations.

Altogether, the complexity is

$$\begin{aligned} T(n) &= \sum_{i=1}^{\lceil \log n \rceil} 2 \log (n/2^i) \\ &= \Theta(\log^2 n). \end{aligned}$$

▷ *Solution 166*

```
FIND-CYCLE(G)
 1  N = array of size |V(G)| // visited
 2  P = array of size |V(G)| // previous
 3  for v ∈ V(G)
 4       N[v] = FALSE
 5       P[v] = NULL
 6  for v ∈ V(G)
 7       if not N[v]
 8            N[v] = TRUE
 9            if FIND-CYCLE-R(N, P, v)
10                 return TRUE
11  return FALSE
```

```
FIND-CYCLE-R(N, P, v)
 1  for w ∈ v.Adj
 2       if N[w]
 3            u = P[v]
 4            while u ≠ NULL
 5                 if u == w
 6                      return TRUE
 7                 u = P[u]
 8       else N[w] = TRUE
 9            P[w] = v
10            if FIND-CYCLE-R(N, P, w)
11                 return TRUE
12  return FALSE
```

▷ *Solution 167.1*

BFS-FIRST-COMMON-ANCESTOR(π, u, v)

```
 1  S = array of size |π|
 2  for i = 1 to |π|
 3      S[i] = 0
 4  while u ≠ NULL or v ≠ NULL
 5      if u ≠ NULL
 6          if S[u] == 1
 7              return u
 8          else S[u] = 1
 9              u = π[u]
10      if v ≠ NULL
11          if S[v] == 1
12              return v
13          else S[v] = 1
14              v = π[v]
15  return NULL
```

The time complexity is $\Theta(n)$. The space complexity is $\Theta(n)$.

▷ *Solution 167.2*

BFS-FIRST-COMMON-ANCESTOR-2(π, D, u, v)

```
1  if D[u] == ∞ or D[v] == ∞
2      return NULL
3  while u ≠ v
4      if D[u] > D[v]
5          u = π[u]
6      else v = π[v]
7  return u
```

The time complexity is $\Theta(n)$.

▷ *Solution 169.1*

BST-FIND-SUM(T, v)

```
1  t1 = BST-MIN(T)
2  while t1 ≠ NULL
3      t2 = BST-SEARCH(T, v − t.key)
4      if t2 ≠ NULL
5          return t1, t2
6      else
7      else t1 = BST-SUCCESSOR(t1)
8  return NULL
```

The time complexity is $\Theta(n^2)$.

▷ *Solution 169.2*

BST-LOWER-BOUND(t, v)

```
    // rightmost element whose key is ≤ v, or NULL
 1  while t ≠ NULL
 2      if v < t.key
 3          t = t.left
 4      elseif t.right ≠ NULL and t.right.key < v
 5          t = t.right
 6      else return t
 7  return NULL
```

BST-FIND-SUM(T, v)

```
 1  t1 = BST-LOWER-BOUND(T, v/2)
 2  t2 = BST-SUCCESSOR(t1)
 3  while t1 ≠ NULL and t2 ≠ NULL
 4      if t1 + t2 = v
 5          return t1, t2
 6      elseif t1 + t2 < v
 7          t2 = BST-SUCCESSOR(t2)
 8      else t1 = BST-PREDECESSOR(t1)
 9  return NULL
```

The time complexity is $\Theta(n)$.

▷ *Solution 170.1*

VERIFY-K-PAIRWISE-RELATIVELY-PRIME(X, k, S)

```
 1  if S ⊈ X or |S| < k
 2      return FALSE
 3  for i = 1 to |S| − 1
 4      for j = i + 1 to |S|
 5          if GCD(S[i], S[j]) > 1
 6              return FALSE
 7  return TRUE
```

GCD(a, b)

```
1  while a ≠ b
2      if a > b
3          a = a % b
4      else b = b % a
5  return a
```

The time complexity is $O(k \log n + k^2 \log m)$, where m is the maximum value in X.

▷ *Solution 177*

function	*rank*
$f_0(n) = n^{n^n}$	1
$f_1(n) = \log^2(n)$	7
$f_2(n) = n!$	2
$f_3(n) = \log(n^2)$	8
$f_4(n) = n$	6
$f_5(n) = \log(n!)$	5
$f_6(n) = \log\log n$	9
$f_7(n) = n \log n$	5
$f_8(n) = \sqrt{n^3}$	4
$f_9(n) = 2^n$	3

▷ *Solution 178*

MINIMAL-COVERING-SQUARE(P)

```
 1  if P.length == 0
 2      return 0
 3  left = P[1].x
 4  right = P[1].x
 5  top = P[1].y
 6  bottom = P[1].y
 7  for i = 2 to P.length
 8      if P[i].x > right
 9          right = P[i].x
10      elseif P[i].x < left
11          left = P[i].x
12      if P[i].y > top
13          top = P[i].y
14      elseif P[i].y < bottom
15          bottom = P[i].y
16  if right − left > top − bottom
17      return (right − left)^2
18  else return (top − bottom)^2
```

▷ *Solution 179*

UNIMODAL-FIND-MAXIMUM(A)

```
 1  l = 1
 2  h = A.length
 3  while l < h − 1
 4      m = ⌊(l + h)/2⌋
 5      if A[m − 1] > A[m]
 6          h = m
 7      elseif A[m + 1] > A[m]
 8          l = m
 9      else return A[m]
10  error "A is not a unimodal sequence"
```

▷ *Solution 180.1*

BETTER-ALGO-X(A, k) returns the beginning and ending position of a minimal subsequence of A that contains at least k equal elements. The complexity of BETTER-ALGO-X is $\Theta(n^4)$. In essence, this is because there are four nested loops.

▷ *Solution 180.2*

Notice that any minimal sequence $P[i], P[i+1], \ldots, P[j]$ that contains at least k equal elements contains *exactly* k elements equal to the first and last element of the sequence. Otherwise, $P[i], \ldots, P[j-1]$ would be a smaller sequence that still contains at least k equal elements.

So, we just have to find a sequence that starts and ends with the same element x, and contains exactly k elements equal to x, including the first and last element:

```
BETTER-ALGO-X(A, k)
 1  l = −∞
 2  r = +∞
 3  for i = 1 to A.length
 4      c = 1
 5      j = i + 1
 6      while c < k and j ≤ min(A.length, i + r − l)
 7          if A[i] == A[j]
 8              c = c + 1
 9          j = j + 1
10      if c == k and r − l > j − i
11          l = i
12          r = j
13  return l, r
```

The complexity of BETTER-ALGO-X is $O(n^2)$.

▷ *Solution 181.1*

```
THREE-WAY-PARTITION(A, begin, end)
 1  q1 = begin
 2  q2 = q1 + 1
 3  for i = q1 + 1 to end − 1
 4      if A[i] ≤ A[q1]
 5          swap A[i] ↔ A[q2]
 6          if A[q2] < A[q1]
 7              swap A[q2] ↔ A[q1]
 8              q1 = q1 + 1
 9          q2 = q2 + 1
10  return q1, q2
```

▷ *Solution 181.2*

```
QUICK-SORT(A)
1  QUICK-SORT-R(A, 1, A.length + 1)
QUICK-SORT-R(A, begin, end)
1  if begin < end
2      q1, q2 = THREE-WAY-PARTITION(A, begin, end)
3      QUICK-SORT-R(A, begin, q1)
4      QUICK-SORT-R(A, q2, end)
```

This variant would be much more efficient with sequences often-repeated elements. In the extreme case of a sequence with n identical numbers, this variant would terminate in time $O(n)$, while the classic algorithm would run in time $O(n^2)$.

▷ *Solution 182*

S(A, s) returns TRUE if there is a subset of the elements in A that add up to s. This is also known as the *subset-sum* problem.

The best-case complexity is $O(n)$. An example of a best-case input (of size n) is with any array A and with $s = 0$. In this case, the algorithm recurses n times in line 3, only then to return TRUE from line 2 of the last recursion, and then unrolling all the recursions out of line 3 to ultimately return TRUE out of line 4.

The worst-case complexity is $O(2^n)$. A worst-case input (of size n) is one that leads to a FALSE result. An example would be an array A of positive numbers with $s < 0$. In this case, every invocation recurses twice, except for the base case. Each recursion reduces the size of the input range by 1, so the recursion tree amounts to a full binary tree with n levels, which leads to a complexity of $O(2^n)$.

▷ *Solution 183*

GROUP-OF-K($S_1, S_2, \ldots, S_m, k$)

```
1  H = empty min-heap (sorted by time)
2  for i = 1 to m
3      t, a = S_i[1]
4      MIN-HEAP-INSERT(H, (t, a, i, 1)) (sorted by t)
5  C = dictionary mapping antennas to integers (hash map)
6  while H is not empty
7      t, a, i, j = MIN-HEAP-EXTRACT-MIN(H)
8      if j > 1
9          t', a' = S_i[j − 1]
10         C[a'] = C[a'] − 1
11     if a ∈ C
12         C[a] = C[a] + 1
13     else C[a] = 1
14     if C[a] ≥ k
15         return t, a
16     if j ≤ S_i.length
17         t, a = S_i[j + 1]
18         MIN-HEAP-INSERT(H, (t, a, i, j + 1)) (sorted by t)
19 return NULL
```

▷ *Solution 184.1*

ALGO-X(A) sorts the elements of A in-place so that all odd numbers precede all even numbers. I other words, ALGO-X(A) *partitions* A in two parts, $A[1 \ldots j-1]$ and $A[j \ldots A.length]$ so that $A[1 \ldots j-1]$ contains only odd numbers and $A[j \ldots A.length]$ contains only even numbers. One of the two parts might be empty. The complexity of ALGO-X is $\Theta(n^2)$.

▷ *Solution 184.2*

BETTER-ALGO-X(A)

```
1  i = 1
2  j = A.length + 1
3  while i < j
4      if A[i] ≡ 0 mod 2 // A[i] is even
5          j = j − 1
6          swap A[i] ↔ A[j]
7      else i = i + 1
8  return j
```

▷ *Solution 185*

BTREE-PRINT-RANGE(T, a, b)

```
1  if not T.leaf and T.key[1] > a
2      BTREE-PRINT-RANGE(T.c[1], a, b)
3  for i = 1 to T.n
4      if T.key[i] ≥ b
5          return
6      if T.key[i] > a
7          print T.key[i]
8      if not T.leaf
9          if i == T.n or T.key[i + 1] > a
10             BTREE-PRINT-RANGE(T.c[i + 1], a, b)
```

▷ *Solution 186*

The problem is in P, and therefore it is also in NP. This is a polynomial-time *solution* algorithm that proves it:

```
ALGO(G, k)
1  for v ∈ V(G)
2      D_v = DIJKSTRA(G, v)
3      // D_v is the distance vector resulting from Dijkstra
4      for u ∈ V(G)
5          if D_v[u] == k
6              return TRUE
7  return FALSE
```

▷ *Solution 187*

```
MOST-CONGESTED-SEGMENT(A, ℓ)
 1  sort A
 2  i = 1
 3  j = 1
 4  x = NULL
 5  m = 0
 6  while j < A.length
 7      if A[j] − A[i] ≤ ℓ
 8          if m < j − i + 1
 9              x = A[i]
10              m = j − i + 1
11          j = j + 1
12      else i = i + 1
13  return x
```

▷ *Solution 188*

The problem is in NP because a TRUE answer can be verified in polynomial time with a "certificate" consisting of a set of nodes $C = \{v_1, v_2, \ldots, v_\ell\}$

```
VERIFY(G, k, C = {v_1, v_2, ..., v_ℓ})
1  if |C| < k
2      return FALSE
3  for all pairs u, v ∈ C
4      if G[u][v] ≠ 1
5          return FALSE
6  return TRUE
```

▷ *Solution 189*

```
MAX-HEAP-TOP-THREE(H)
 1  if H.length < 4
 2      for i = 1 to H.length
 3  else print(H[1])
 4      if H[2] > H[3]
 5          i = 2
 6          j = 3
 7      else i = 3
 8          j = 2
 9      if H.length ≥ 2i + 1 and H[j] < H[2i + 1]
10          j = 2i + 1
11      if H.length ≥ 2i and H[j] < H[2i]
12          j = 2i
13      print(H[i])
14      print(H[j])
```

▷ *Solution 190*

```
LONGEST-STRETCH(P, h)
 1  ℓ = 0
 2  i = 1
 3  while i < P.length
 4       a = P[i].y
 5       b = P[i].y
 6       j = i + 1
 7       while b − a < h
 8            if P[j].y > b
 9                 b = P[j].y
10            elseif P[j].y < a
11                 a = P[j].y
12            if b − a < h
13                 if P[j].x − P[i].x > ℓ
14                      ℓ = P[j].x − P[i].x
15            else j = j + 1
16       i = i + 1
17  return ℓ
```

LONGEST-STRETCH(P, h) runs in $O(n^2)$ in the worst case. For example, a completely flat road would be a worst-case input.

▷ *Solution 191*

```
IS-BIPARTITE(G)
 1  for v ∈ V(G)
 2       C[v] = GREEN// can be in either V_A or V_B
 3  for v ∈ V(G)
 4  if C[v] == GREEN
 5       C[v] = RED
 6       Q = {v}// queue containing v
 7       while Q is not empty
 8            u = DEQUEUE(Q)
 9            for all w adjacent to u:
10                 if C[w] == GREEN
11                      if C[v] == RED
12                           C[w] = BLUE
13                      else C[w] = RED
14                      ENQUEUE(Q, w)
15                 else if C[v] ≠ C[w]
16                           return FALSE
17  return TRUE
```

▷ *Solution 192.1*

```
GOOD-ARE-ADJACENT(A)
1  i = 1
2  while i < j and not IS-GOOD(A[i])
3       i = i + 1
4  while i < j and not IS-GOOD(A[j])
5       j = j − 1
6  while i < j
7       if not IS-GOOD(A[i])
8       else i = i + 1
9  return TRUE
```

▷ *Solution 192.2*

MAKE-GOOD-ADJACENT(A)

```
1  i = 1
2  while i < j and not IS-GOOD(A[i])
3        i = i + 1
4  while i < j and not IS-GOOD(A[j])
5        j = j − 1
6  while i < j
7        if not IS-GOOD(A[i])
8              swap A[i] ↔ A[j]
9              j = j − 1
10       i = i + 1
11 return TRUE
```

▷ *Solution 193*

The problem is in P, and therefore also in NP. This is an algorithm that solves the problem in $O(n \log n)$ time.

GROUP-OF-EQUALS(A, k)

```
1  B = SORT(A)
2  i = 1
3  j = 1
4  while j < A.length
5        if A[i] == A[j]
6              j = j + 1
7              if j − i == k
8                    return TRUE
9        else i = j
10       return FALSE
```

▷ *Solution 194*

A simple, brute-force solution is to check each combination of positions in the two strings

MAXIMAL-COMMON-SUBSTRING(X, Y)

```
1  m = 0
2  for i = 1 to A.length
3        for j = 1 to B.length
4              ℓ = 0
5              while i + ℓ ≤ A.length and j + ℓ ≤ B.length and A[i + ℓ] == B[j + ℓ]
6                    ℓ = ℓ + 1
7              if ℓ > m
8                    m = ℓ
9  return m
```

The complexity of MAXIMAL-COMMON-SUBSTRING is $O(n^3)$.

▷ *Solution 195*

BST-COUNT-UNBALANCED-NODES(t)

```
1  if t == NULL
2        return 0, 0
3  U_L, Tot_L = BST-COUNT-UNBALANCED-NODES(t.left)
4  U_R, Tot_R = BST-COUNT-UNBALANCED-NODES(t.right)
5  U = U_L + U_R
6  if Tot_L > 2Tot_R + 1 or Tot_R > 2Tot_L + 1
7        U = U + 1
8  return U, (Tot_L + Tot_R + 1)
```

The complexity is $\Theta(n)$.

▷ *Solution 196.1*
ALGO-X returns the maximal difference between two values in an increasing sequence of elements in A. The complexity is $\Theta(n^3)$.

▷ *Solution 196.2*

```
LINEAR-ALGO-X(A)
 1  x = 0
 2  i = 0
 3  j = 1
 4  while j ≤ A.length
 5       if A[j] > A[j − 1]
 6            if A[j] − A[i] > x
 7                 x = A[j] − A[i]
 8       else i = j
 9       j = j + 1
10  return x
```

▷ *Solution 197.1*
A naïve solution for FIND-SQUARE is to test all quadruples of points p_i, p_j, p_k, p_ℓ, and determine whether p_i, p_j, p_k, p_ℓ form a square.

```
FIND-SQUARE(P)
 1  for i = 1 to P.length
 2       for j = 1 to P.length
 3            for k = 1 to P.length
 4                 for ℓ = 1 to P.length
 5                      d_x = P[j].x − P[i].x
 6                      d_y = P[j].y − P[i].y
 7                      if P[k].x == P[j].x + d_y and P[k].y == P[j].y − d_x
                        and P[ℓ].x == P[i].x + d_y and P[ℓ].y == P[i].y − d_x
 8                           return TRUE
 9  return FALSE
```

▷ *Solution 197.2*
Here the idea is to test all segments defined by two distinct points, and then to try to find the other corners of a square, which we can do with a binary search.

```
ORDER-2D(p_1, p_2)
 1  if p_1.x < p_2.x
 2       return TRUE
 3  elseif p_1.x > p_2.x
 4       return FALSE
 5  elseif p_1.y < p_2.y
 6       return TRUE
 7  elseif p_1.y < p_2.y
 8  else return FALSE
```

```
BINARY-SEARCH-2D(P, x, y)
 1  i = 1
 2  j = P.length
 3  while i ≤ j
 4       m = ⌊(i + j)/2⌋
 5       if ORDER-2D(v, P[m])
 6            j = m − 1
 7       elseif P[m].x == x and P[m].y == y
 8            return TRUE
 9       else i = m + 1
10  return FALSE
```

```
FIND-SQUARE(P)
1  sort P using ORDER-2D as a comparison between pairs of points
2  for i = 1 to P.length
3      for j = 1 to P.length
4          d_x = P[j].x − P[i].x
5          d_y = P[j].y − P[i].y
6          if BINARY-SEARCH-2D(P, P[i].x + d_y, P[i].y − d_x)
           and BINARY-SEARCH-2D(P, P[j].x + d_y, P[j].y − d_x)
7              return TRUE
8  return FALSE
```

▷ *Solution 198*

```
INITIALIZE(Q)
1  Q.A = new empty array
2  Q.P = new empty array
```

```
ENQUEUE(Q, obj, p)
1  append obj to array Q.A
2  append p to array Q.P
3  i = Q.P.length
4  j = ⌊i/2⌋
5  while i > 1 and Q.P[i] < Q.P[j]
6      swap Q.P[i] ↔ Q.P[j]
7      swap Q.A[i] ↔ Q.A[j]
8      i = j
9      j = ⌊i/2⌋
```

```
DEQUEUE(Q)
 1  ℓ = A.P.length
 2  if ℓ < 1
 3      error "empty queue"
 4  x = Q.A[1]
 5  swap Q.P[1] = Q.P[ℓ]
 6  swap Q.A[1] = Q.A[ℓ]
 7  remove last element from Q.P
 8  remove last element from Q.A
 9  ℓ = ℓ − 1
10  i = 1
11  while 2i ≤ ℓ and Q.P[i] > Q.P[2i]
    or 2i + 1 ≤ ℓ and Q.P[i] > Q.P[2i + 1]
12      if 2i + 1 ≤ ℓ and Q.P[2i + 1] > Q.P[2i]
13          j = 2i + 1
14      else j = 2i
15          swap Q.P[i] = Q.P[j]
16          swap Q.A[i] = Q.A[j]
17          i = j
18  return x
```

▷ *Solution 199*

```
MAXIMAL-DISTANCE(A)
 1  if A.length < 2
 2      return 0
 3  min = A[1]
 4  max = A[1]
 5  for i = 2 to A.length
 6      if A[i] > max
 7          max = A[i]
 8      elseif A[i] < min
 9          min = A[i]
10  return max − min
```

▷ *Solution 200*

BST-HEIGHT(t)

```
1  if t == NULL
2      return 0
3  return 1 + max(BST-HEIGHT(t.left), BST-HEIGHT(t.right))
```

▷ *Solution 201*

The problem, which is the well-known *matching* problem in graph theory, is definitely in NP. This is a possible verification algorithm:

VERIFY-MATCHING($G = (V, E, w), t, S$)

```
1  X = ∅
2  for e = (u, v) ∈ S
3      if u ∈ X or v ∈ X
4          return FALSE
5      X = X ∪ {u, v}
6      weight = weight + w(e)
7  if weight ≥ t
8      return TRUE
9  else return FALSE
```

▷ *Solution 202*

We can use a dynamic programming approach. Let P_i be the maximal value of the objects you can collect by reaching object i. Now, since you can reach P_i only by increasing your x and y coordinates, then that means that the maximal total value P_i is the value of object i plus the maximal total value when you reach any one of the objects from which you can then reach object i. This means all the objects with coordinates less than those of i. So:

$$P_i = V[i] + \max_{j|X[j]\leq X[i]\wedge Y[j]\leq Y[i]} P[j]$$

The global maximal game value is then $\max P_i$.

Now, the formula for P_i gives us a very simple recursive algorithm. This is inefficient, but it can be made very efficient with memoization.

MAXIMAL-GAME-VALUE(X, Y, V)

```
1  P = array of n = |V| elements initialized to P[i] = NULL
2  m = −∞
3  for i = 1 to V.length
4      if m < MAXIMAL-VALUE-P(P, X, Y, V, i)
5          m = MAXIMAL-VALUE-P(P, X, Y, V, i)
6  return m
```

MAXIMAL-VALUE-P(P, X, Y, V, i)

```
1  if P[i] == NULL
2      P[i] = V[i]
3      for j = 1 to V.length
4          if j ≠ i and X[j] ≤ X[i] and Y[j] ≤ Y[i]
5              if P[i] < V[i] + MAXIMAL-VALUE-P(P, X, Y, V, j)
6                  P[i] = V[i] + MAXIMAL-VALUE-P(P, X, Y, V, j)
7  return P[i]
```

▷ *Solution 203*

We are not required to be particularly efficient, so we can write a simple algorithm.

```
MAXIMAL-SUBSTRING(S)
 1  A = NULL
 2  for i = 1 to |S|
 3       X = ∅
 4       for j = 1 to |S[i]| − 1
 5            for k = j + 1 to |S[i]|
 6                 X = X ∪ {S[i][j ... k]}
 7       if A == NULL
 8            A = X
 9       else A = A ∩ X
10            if A == ∅
11                 return “”
12  return longest string in A or “” if A == NULL
```

▷ *Solution 204.1*

ALGO-X returns *mode* of A, meaning an element that occurs in A with maximal frequency (count). The complexity is $\Theta(n^2)$. Any input is the worst-case input.

▷ *Solution 204.2*

```
BETTER-ALGO-X(A)
 1  B = copy of A
 2  sort B
 3  if |S| == 0
 4       return 0
 5  x = B[1]
 6  m = 1
 7  c = 1
 8  for i = 2 to |S|
 9       if B[i] == B[i − 1]
10            c = c + 1
11            if c > m
12                 m = c
13                 x = B[i]
14       else c = 1
15  return x
```

▷ *Solution 205*

```
GRAPH-DEGREE(G)
 1  n = |V(G)|
 2  for i = 1 to n
 3       d = 0
 4       for j = 1 to n
 5            if G[i, j] == 1
 6                 d = d + 1
 7       if d > m
 8            m = d
 9  return m
```

▷ *Solution 206*

We don’t have complexity constraints, so the algorithm can be simple:

```
FIND-3-CYCLE(G)
1  for u ∈ V(G)
2      for v ∈ Adj[u]
3          for w ∈ Adj[v]
4              for x ∈ Adj[w]
5                  if x == u
6                      return TRUE
7  return FALSE
```

The complexity is $\Theta(n\Delta^3)$, where Δ is the degree. Now, consider the full bipartite graph of $n/2$ plus $n/2$ nodes. In this case, the complexity is $\Theta(n^4)$.

▷ *Solution 207*

Here's an obvious $O(mn^2)$ solution:

```
LONGEST-COMMON-PREFIX(S)
1  m = 0
2  for i = 2 to S.length
3      for j = 1 to i − 1
4          ℓ = COMMON-PREFIX-LENGTH(S[i], S[j])
5          if ℓ > m
6              ℓ = m
7  return m
```

```
COMMON-PREFIX-LENGTH(a, b)
1  for i = 1 to min(a.length, b.length)
2      if a[i] ≠ b[i]
3          return i − 1
4  return min(a.length, b.length)
```

▷ *Solution 208*

Notice that if we sort the array S in lexicographical order, then any k strings with a common prefix will be contiguous in the sorted array.

```
LONGEST-K-COMMON-PREFIX(S, k)
1  sort S in lexicographical order
2  m = 0
3  for i = k to S.length
4      ℓ = COMMON-PREFIX-LENGTH(S[i], S[i − k])
5      if ℓ > m
6          ℓ = m
7  return m
```

```
COMMON-PREFIX-LENGTH(a, b)
1  for i = 1 to min(a.length, b.length)
2      if a[i] ≠ b[i]
3          return i − 1
4  return min(a.length, b.length)
```

This complexity is $O(m \log n)$.

▷ *Solution 209*

The problem is in P and therefore it is also in NP. This is an algorithm that solves the problem in polynomial time:

```
NEGATIVE-THREE-CYCLE(G = (V, E))
1  for u ∈ V
2      for v ∈ Adj[u]
3          for w ∈ Adj[v]
4              if w == u and weight(u, v) + weight(v, w) + weight(w, u) < 0
5                  return TRUE
6  return FALSE
```

▷ *Solution 210.1*

```
UPPER-BOUND(A, x)
 1  u = UNDEFINED
 2  d = UNDEFINED
 3  for a ∈ A
 4       if x ≤ a
 5            if u == UNDEFINED or d > a − x
 6                 d = a − x
 7                 u = a
 8  return u
```

The complexity is $\Theta(n)$.

▷ *Solution 210.2*

```
UPPER-BOUND-SORTED(A, x)
 1  i = 1
 2  j = A.length
 3  if A[j] < x
 4       return UNDEFINED
 5  elseif A[i] ≥ x
 6       return A[i]
 7  while i < j
 8       m = ⌊i + j/2⌋
 9       if A[m] == x
10            return A[m]
11       elseif A[m] < x
12            i = m
13       else j = m
14  return A[j]
```

The complexity is $\Theta(\log n)$.

▷ *Solution 210.3*

```
UPPER-BOUND-BST(T, x)
 1  while T ≠ NULL
 2       if T.key < x
 3            T = T.right
 4       else while T.left ≠ NULL and T.left.key ≥ x
 5                 T = T.left
 6            return T.key
 7  return UNDEFINED
```

The complexity is $\Theta(h)$ where h is the height of the input tree.

▷ *Solution 211*

```
SUM-OF-THREE(A, s)
 1  B = MERGE-SORT(A)
 2  for i = 1 to A.length
 3       j = 1
 4       k = A.length
 5       while j < k
 6            if j == i
 7                 j = j + 1
 8            elseif k == i
 9                 k = k − 1
10            elseif B[i] + B[j] + B[k] == s
11                 return TRUE
12            elseif B[i] + B[j] + B[k] > s
13                 k = k − 1
14            else j = j + 1
15  return FALSE
```

▷ *Solution 212.1*

ALGO-X returns TRUE if and only if A contains two distinct numbers whose absolute difference is greater than x.

The complexity of ALGO-X is $\Theta(n^2)$. The worst case is when the algorithm reaches the last return statement in line 10. In this case, the loop amounts to an iteration over all pairs of distinct positions $j < i$. In fact, the loop starts with j and i at beginning and end of the array, respectively. Then the loop moves j forward by one step at a time until j reaches i, at which point j restarts from the beginning and i moves by one position to the left.

▷ *Solution 212.2*

```
BETTER-ALGO-X(A, x)
 1  if A.length < 1
 2       return FALSE
 3  low = A[1]
 4  high = A[1]
 5  for i = 2 to A.length
 6       if A[i] < low
 7            low = A[i]
 8       elseif A[i] > high
 9            high = A[i]
10       if high − low > x
11            return TRUE
12  return FALSE
```

BETTER-ALGO-X scans the array once looking for the maximum and minimum values, and exits as soon as it finds that the difference between the current (partial) maximum and minimum is greater than x.

▷ *Solution 213.1*

ALGO-S returns the value x in A such that there are exactly k elements less than x in A, or NULL if no such element exists.

The complexity of ALGO-S is $\Theta(n^2)$. The worst case is when the algorithm returns NULL. In this case, ALGO-S loops for exactly n times, each one costing the complexity of ALGO-R running on A, which is also n iterations.

▷ *Solution 213.2*

```
BETTER-ALGO-S(A, k)
 1  B = MERGE-SORT(A)
 2  if k ≥ 0 and k < B.length and B[k + 1] > B[k]
 3        return B[k + 1]
 4  else return NULL
```

▷ *Solution 214*

```
SORT-SPECIAL(A)
 1  q = A.length
 2  while q > 1
 3        for i = 1 to q − 1
 4              if A[i] > A[q]
 5                    swap A[i] ↔ A[q]
 6        i = 1
 7        while i < q
 8              if A[i] == A[q]
 9                    swap A[i] ↔ A[q − 1]
10                    q = q − 1
11              else i = i + 1
```

▷ *Solution 215*

```
HEAP-PROPERTIES(A)
 1  max = 1
 2  min = 1
 3  for i = A.length downto 2
 4        p = ⌊i/2⌋
 5        if A[i] < A[p]
 6              min = 0
 7        elseif A[i] > A[p]
 8              max = 0
 9  if min == 1
10        if max == 1
11              return 2
12        else return −1
13  else if max == 1
14              return 1
15        else return 0
```

▷ *Solution 216*

We can find and then group, and thereby count, pairs of compatible objects as follows:

```
MAX-COMPATIBLE-PAIRING(A)
 1  c = 0
 2  i = 1
 3  while i < A.length
 4        j = i + 1
 5        while j ≤ A.length and COMPATIBLE(A[i], A[j]) == FALSE
 6              j = j + 1
 7        if j ≤ A.length // we found a compatible pair (i, j)
 8              swap A[i + 1] ↔ A[j]
 9              i = i + 2
10              c = c + 1
11        else i = i + 1
12  return c
```

Another option is not to group but rather to simply count the pairs of equivalent elements. However, to avoid counting elements multiple times, we must somehow mark elements as being part of pair.

```
MAX-COMPATIBLE-PAIRING(A)
 1  c = 0
 2  B = [0, 0, ..., 0] // array of n Boolean values
 3  i = 1
 4  while i < A.length
 5      j = i + 1
 6      while j ≤ A.length and (B[j] == 1 or COMPATIBLE(A[i], A[j]) == FALSE)
 7          j = i + 1
 8      if j ≤ A.length // we found a compatible pair (i, j)
 9          B[j] = 1
10          i = i + 1
11          c = c + 1
12      else i = i + 1
13  return c
```

▷ *Solution 217*

A queen in row i and column j attacks all the squares in row x and column y such that $x = i$, $y = j$, $x + y = i + j$, or $x - y = i - j$. Therefore, we can simply create an index for each of these four conditions for all the white queens, and then check whether a black queen is in any one of these indexes. The following solution creates indexes using sorted arrays. Another solution would be to use hash tables.

```
WHITE-ATTACKS-BLACK(W, B)
 1  R = array of W.length integers
 2  C = array of W.length integers
 3  D1 = array of W.length integers
 4  D2 = array of W.length integers
 5  for i = 1 to W.length
 6      R[i] = W[i].row
 7      C[i] = W[i].col
 8      D1[i] = W[i].row − W[i].col
 9      D2[i] = W[i].row + W[i].col
10  sort R
11  sort C
12  sort D1
13  sort D2
14  for i = 1 to B.length
15      if BINARY-SEARCH(R, B[i].row)
16          return TRUE
17      if BINARY-SEARCH(C, B[i].col)
18          return TRUE
19      if BINARY-SEARCH(D1, B[i].row − B[i].col)
20          return TRUE
21      if BINARY-SEARCH(D2, B[i].row + B[i].col)
22          return TRUE
23  return FALSE
```

The complexity is $O(n \log n)$, since the sorting phase for the white queens takes $O(n \log n)$, and the lookup phase for the black queens also takes $O(n \log n)$.

▷ *Solution 218.1*

A very simple solution is the following recursive algorithm:

```
COUNT-FULL-NODES(t)
 1  if t == NULL
 2      return 0
 3  if t.left ≠ NULL and t.right ≠ NULL
 4      return 1 + COUNT-FULL-NODES(t.left) + COUNT-FULL-NODES(t.right)
 5  else return COUNT-FULL-NODES(t.left) + COUNT-FULL-NODES(t.right)
```

The algorithm is equivalent to a walk of the tree, so it visits each node exactly once. The complexity is therefore $\Theta(n)$.

▷ *Solution 218.2*

The idea here is to walk throgh the nodes that are not full, and to rotate those that are full until they have a single child. For example, right rotate until they have only a right child. This leads to the following recursive solution.

```
NO-FULL-NODES(t)
 1  if t == NULL
 2      return t
 3  if t.left == NULL
 4      t.right = NO-FULL-NODES(t.right)
 5      return t
 6  elseif t.right == NULL
 7      t.left = NO-FULL-NODES(t.left)
 8      return t
 9  while t.left ≠ NULL
10      t = RIGHT-ROTATE(t)
11  t.right = NO-FULL-NODES(t.right)
12  return t
```

The algorithm is also a tree walk for all non-full nodes. For each node t that is full, the algorithm performs some right rotations so as to move all the nodes in the subtree rooted at t from the left to the right side of t. This process iterates through those nodes at most once. So, in total, each node is touched either once or twice by the algorithm. The complexity is therefore $\Theta(n)$.

▷ *Solution 219.1*

The problem is in NP, since given a permutation Π of the indexes, it is easy to show (in polynomial time) that the answer is indeed TRUE with the following algorithm:

```
VERIFY-PERM-SUM-K(A, B, Π)
 1  if A.length ≠ B.length
 2      return FALSE
 3  n = A.length
 4  A' = array of n NULL values
 5  for i = 1 to n
 6      A'[Π(i)] = A[i]
 7  for i = 1 to n
 8      if A'[i] == NULL
 9          return FALSE
10  k = A'[1] + B[1]
11  for i = 2 to n
12      if A'[i] + B[i] ≠ k
13          return FALSE
14  return TRUE
```

This algorithm first checks that Π indeed defines a permutation on A, then it checks that the permutation satisfies the condition of the problem.

This question can also be immediately answered by solving the following exercise question, that is, showing that the problem is in P.

▷ *Solution 219.2*

The problem is in P, since it is easy to show that if A' exists, then it is also possible to sort A in increasing order, and correspondingly B in decreasing order so that the condition is satisfied. So, the following algorithm *solves* the problem, in polynomial time.

```
PERM-SUM-K(A, B)
1  if A.length ≠ B.length
2       return FALSE
3  A' = sort A
4  B' = sort B in reverse order
5  k = A'[1] + B'[1]
6  for i = 2 to n
7       if A'[i] + B'[i] ≠ k
8            return FALSE
9  return TRUE
```

▷ *Solution 220*

There is a simple dynamic-programming solution. Let m_i be the minimal contiguous sub-sequence sum *ending* at position i. Then, we can obtain the minimal contiguous subsequence ending at $i+1$ by either connecting to the minimal contiguous subsequence ending at i, or by starting and ending a singleton subsequence at $i+1$. So, $m_{i+1} = \min\{m_i + A[i], A[i]\}$. In the case of the first element $A[1]$, the value of m_1 is simply $A[1]$. Then, from there we can compute all the other values, and remember the minimal value.

```
MINIMAL-CONTIGUOUS-SUM(A)
1  m = A[1]
2  p = A[1]
3  for i = 2 to A.length
4       if p > 0
5            p = A[i]
6       else p = p + A[i]
7       if p < m
8            m = p
9  return m
```

▷ *Solution 221*

The idea is to find whether there is any path that can turn onto itself. We can do that using a simple depth-first search. We mark a node v as *visited* when we find v for the first time, and then we mark v as *finished* when we have explored all the nodes reachable from v. We have a cycle when, from the current node v we hit a neighbor u that is visited but not yet finished. This is because this means that v is reachable from u, and v is reachable from u (it is one of its neighbors), so we have a cycle.

```
HAS-CYCLE(G)
1  visited = ∅
2  finished = ∅
3  for all v ∈ V(G)
4       if DFS-FIND-CYCLE(G, v)
5            return TRUE
6  return FALSE
```

```
DFS-FIND-CYCLE(G, v)
1  global variable visited
2  global variable finished
3  if v ∈ visited
4       if v ∉ finished
5            return TRUE
6       else return FALSE
7  visited = visited ∪ {v}
8  for all u ∈ G.Adj[v]
9       if DFS-FIND-CYCLE(G, u)
10           return TRUE
11 finished = finished ∪ {v}
12 return FALSE
```

▷ *Solution 222*

A DNA sequence S_1 is a permutation of another DNA sequence S_2 when S_1 contains exactly the same number of A's, the same number of C's, the same number of G's, and the same number of T's as S_2. So, it is easy to check that S_1 is a permutation of S_2 by counting and comparing the numbers of A's, C's, G's, and T's, respectively, in both sequences.
So, let $m = |X|$ be the length of the X sequence, then a basic idea would be to look at each substring $S[i, \ldots, i + m - 1]$ in S, and check whether $S[i, \ldots, i + m - 1]$ is a permutation of X.

DNA-PERMUTATION-SUBSTRING(S, X)

```
 1  x_A = number of A's in X
 2  x_C = number of C's in X
 3  x_G = number of G's in X
 4  x_T = number of T's in X
 5  for i = 1 to S.length − X.length + 1
 6      s_A = number of A's in S[i, ..., i + X.length − 1]
 7      s_C = number of C's in S[i, ..., i + X.length − 1]
 8      s_G = number of G's in S[i, ..., i + X.length − 1]
 9      s_T = number of T's in S[i, ..., i + X.length − 1]
10      if x_A == s_A and x_C == s_C and x_G == s_G and x_T == s_T
11          return TRUE
12  return FALSE
```

However, the complexity of this algorithm is $\Theta(\ell m)$, where $\ell = |S|$ and $m = |X|$. Instead, we want $O(n)$ where $n = \ell + m$.
We can do this with the following algorithm:

DNA-PERMUTATION-SUBSTRING(S, X)

```
 1  // the following can be computed in time linear in |X|
 2  x_A = number of A's in X
 3  x_C = number of C's in X
 4  x_G = number of G's in X
 5  x_T = number of T's in X
 6  // the following can be computed in time linear in |S|
 7  let A, C, G, T be arrays of length n + 1
    such that A[i] is the number of A's in the first i − 1 symbols of S
    and correspondingly for C, G, and T arrays
 8  for i = m + 1 to n + 1
 9      if x_A == A[i] − A[i − m] and x_C == C[i] − C[i − m]
        and x_G == G[i] − G[i − m] and x_T == T[i] − T[i − m]
10          return TRUE
11  return FALSE
```

More in detail:

```
DNA-PERMUTATION-SUBSTRING(S, X)
 1  n = S.length
 2  m = X.length
 3  x_A = 0
 4  x_C = 0
 5  x_G = 0
 6  x_T = 0
 7  for i = 1 to m
 8       if X[i] == 'A'
 9            x_a = x_a + 1
10       elseif X[i] == 'C'
11            x_C = x_C + 1
12       elseif X[i] == 'G'
13            x_G = x_G + 1
14       elseif X[i] == 'T'
15            x_T = x_T + 1
16  let A, C, G, T be arrays of length S.length + 1
17  A[1] = 0
18  C[1] = 0
19  G[1] = 0
20  T[1] = 0
21  for i = 2 to S.length + 1
22       A[i] = A[i − 1]
23       C[i] = C[i − 1]
24       G[i] = G[i − 1]
25       T[i] = T[i − 1]
26       if S[i − 1] == 'A'
27            A[i] = A[i] + 1
28       elseif S[i − 1] == 'C'
29            C[i] = C[i] + 1
30       elseif S[i − 1] == 'G'
31            G[i] = G[i] + 1
32       elseif S[i − 1] == 'T'
33            T[i] = T[i] + 1
34  for i = m + 1 to S.length + 1
35       if x_A == A[i] − A[i − m] and x_C == C[i] − C[i − m]
         and x_G == G[i] − G[i − m] and x_T == T[i] − T[i − m]
36            return TRUE
37  return FALSE
```

▷ *Solution 223.1*

ALGO-X returns the lowest, most common number in A. That is, the number x such that no other number appears more often than x in A, and if there is other number y that appear exactly the same number of times as x, then $x < y$. The complexity of ALGO-X is $\Theta(n^2)$.

▷ *Solution 223.2*

BETTER-ALGO-X(A)

```
1  B = copy of A sorted in ascending order
2  x = B[1]
3  m = 1
4  i = 2
5  while i < A.length
6       j = i + 1
7       while B[i] == B[j]
8            j = j + 1
9       if j − i > m
10           m = j − 1
11           x = B[i]
12      i = j
13 return x
```

The complexity of BETTER-ALGO-X is $\Theta(n \log n)$, since the loop amounts to a linear scan of the sorted array B, so the dominating complexity is the time needed to sort the array at the beginning, which can be done in $\Theta(n \log n)$.

▷ *Solution 224.1*

BST-EQUALS(t_1, t_2)

```
1  if t1 == NIL and t2 == NIL
2       return TRUE
3  elseif t1 == NIL or t2 == NIL
4       return FALSE
5  if t1.key == t2.key and BST-EQUALS(t1.left, t2.left) and BST-EQUALS(t1.right, t2.right)
6       return TRUE
7  else return FALSE
```

The algorithm amounts to a parallel walk of both trees. The complexity is therefore $O(n)$.

▷ *Solution 224.2*
We iterate through all the keys *in order* in both trees, and check them one by one.

BST-MIN-NODE(t)

```
1  if t == NIL
2       return NIL
3  while t.left ≠ NIL
4       t = t.left
5  return t
```

BST-SUCCESSOR(t)

```
1  if t == NIL
2       return NIL
3  elseif t.right ≠ NIL
4       return BST-MIN-NODE(t.right
5  else while t.parent ≠ NIL
6            if t == t.parentleft
7                 return t.parent
8            else t = t.parent
9       return NIL
```

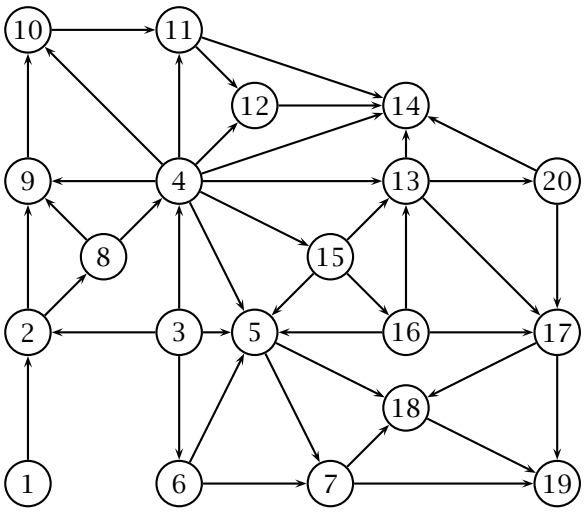
```
BST-EQUAL-KEYS(t_1, t_2)
 1  t_1 = BST-MIN-NODE(t_1)
 2  t_2 = BST-MIN-NODE(t_2)
 3  while t_1 ≠ NIL and t_2 ≠ NIL
 4        if t_1.key ≠ t_2.key
 5              return FALSE
 6        t_1 = BST-SUCCESSOR(t_1)
 7        t_2 = BST-SUCCESSOR(t_2)
 8  if t_1 == NIL and t_2 == NIL
 9        return TRUE
10  else return FALSE
```

The algorithm amounts to a parallel walk of both trees. The complexity is therefore $O(n)$.

▷ *Solution 225*

```
KNIGHT-DISTANCE(r_1, c_1, r_2, c_2)
 1  return KNIGHT-DISTANCE-ORIGIN(r_1 − r_2, c_1 − c_2)
KNIGHT-DISTANCE-ORIGIN(x, y)
 1  M = global "memoization" matrix/dictionary initialized as follows:
    M[0,0] = 0
    M[0,1] = 3
    M[0,2] = 2
    M[1,1] = 2
    M[1,2] = 1
    M[2,2] = 4
    any other value in M is initially NULL
 2  if x < 0
 3        x = −x
 4  if y < 0
 5        y = −y
 6  if x > y
 7        swap x and y
 8  if M[x,y] ≠ NULL
 9        return M[x,y]
10  d_1 = KNIGHT-DISTANCE-ORIGIN(x − 2, y − 1) + 1
11  d_2 = KNIGHT-DISTANCE-ORIGIN(x − 1, y − 2) + 1
12  M[x,y] = min(d1, d2)
13  return min(d1, d2)
```

▷ *Solution 226.1*



▷ *Solution 229.1*

ALGO-X sorts the input array in-place. In the best case, the algorithm terminates in the first execution of the outer loop, with the condition s == TRUE. This is the case when the inner loop does not swap a single element of the array, meaning that the array is already sorted. So, the best-case complexity is $O(n)$. Conversely, the worst case is when each iteration of the outer loop swaps at least one element. This happens when the array is sorted in reverse order. So, the worst-case complexity is $O(n^2)$.

▷ *Solution 229.2*

ALGO-Y sorts the input array in-place so that the value $v = A[0]$, that is the element originally at position 0, ends up in position q, and every other element less than v ends up somewhere in $A[1 \ldots q-1]$, that is to the left of q, and every other element less than or equal to v ends up somewhere in $A[q+1 \ldots |A|]$. In other words, ALGO-Y partitions the input array in-place using the first element as the "pivot". The loop closes the gap between i and j, which are initially the first and last position in the array, respectively. Each iteration either moves i to the right or j to the left, so each iteration reduces the gap by one. Therefore, in any case—worst case is the same as the best case—the complexity is $O(n)$.

▷ *Solution 230*

```
PARTITION-ZERO(A)
 1  j = 1
 2  for i = 1 to A.length
 3       if A[i] < 0
 4            swap A[i] ↔ A[j]
 5            j = j + 1
 6  for i = j to A.length
 7       if A[i] == 0
 8            swap A[i] ↔ A[j]
 9            j = j + 1
```

▷ *Solution 231*

We can use a binary heap to implement a priority queue. In particular, we can use a max-heap where the heap property is based on the comparison between object priorities. Therefore we use an array Q as the base data structure, and we also keep track of the queue size $Q.size$, meaning the number of elements in the queue. Note that this is not the allocated size of the array $Q.size$.

```
PQ-INIT(n)
1  Q = new array of size n
2  Q.size = 0
3  Q.maxsize = n
4  return Q
```

```
PQ-ENQUEUE(Q, x)
1  if Q.size == Q.maxsize
2       return "error: queue overflow"
3  Q[Q.size] = x
4  i = Q.size
5  Q.size = Q.size + 1
6  while i > 1 and Q[i] > Q[⌊i/2⌋]
7       swap Q[i] ↔ Q[⌊i/2⌋]n
8       i = ⌊i/2⌋
```

```
PQ-DEQUEUE(Q)
 1  if Q.size == 0
 2       return "error: empty queue"
 3  x = Q[1]
 4  swap Q[1] ↔ Q[Q.size]
 5  Q.size = Q.size + 1
 6  i = 1
 7  while 2i ≤ Q.size
 8       j = i
 9       m = Q[i]
10       if Q[2i] > m
11            j = 2i
12            m = Q[2i]
13       if 2i + 1 ≤ Q.size and Q[2i + 1] > m
14            j = 2i + 1
15            m = Q[2i + 1]
16       if j > i
17            swap Q[i] ↔ Q[j]
18            i = j
19       else return x
20  return x
```

▷ *Solution 232.1*
ALGO-X returns TRUE if and only if B contains a subset of the elements in A. The complexity is $\Theta(n^2)$. A worst case input is one in which $A.length = B.length = n/2$ and none of the elements of A are contained in B. In this case, the outer loop (line 3) runs for $n/2$ iterations, and the inner loop also runs for $n/2$ times for each of the iterations of the outer loop.

▷ *Solution 232.2*
A different strategy is to first sort both vectors, and then to use an algorithm similar to a merge as below.

```
BETTER-ALGO-X(A, B)
 1  C = sorted copy of array A
 2  D = sorted copy of array B
 3  n = D.length
 4  j = 1
 5  i = 1
 6  while i ≤ C.length and j ≤ D.length
 7        if C[i] < D[j]
 8              i = i + 1
 9        elseif D[i] > D[j]
10              j = j + 1
11        else n = n − 1
12              j = j + 1
13              i = i + 1
14  if n == 0
15        return TRUE
16  else return FALSE
```

▷ *Solution 233.1*
QUESTIONABLE-SORT is correct. It is also known as *selection-sort*. Effectively, the inner loop (j-loop) finds and leaves in position i a *minimal* element in $A[i+1 \ldots n]$. ALGO-X returns TRUE if and only if B contains a subset of the elements in A. The complexity is $\Theta(n^2)$. A worst case input is one in which $A.length = B.length = n/2$ and none of the elements of A are contained in B. In this case, the outer loop (line 3) runs for $n/2$ iterations, and the inner loop also runs for $n/2$ times for each of the iterations of the outer loop.

▷ *Solution 233.2*
Quick-sort or heap-sort would work. Here's quick-sort:

```
BETTER-SORT(A)
 1  QUICK-SORT(A, 1, A.length)

QUICK-SORT(A, b, e)
 1  if e − b > 0
 2        q = PARTITION(A, b, e)
 3        QUICK-SORT(A, b, q − 1)
 4        QUICK-SORT(A, q + 1, e)
```

```
PARTITION(A, b, e)
 1  q = random position in [b, ..., e]
 2  swap A[q] ↔ A[e]
 3  q = b
 4  i = b
 5  for i = b to e
 6        if A[i] ≤ A[e]
 7              swap A[q] ↔ A[i]
 8              q = q + 1
 9  return q − 1
```

▷ *Solution 234*
We can use a binary search.

```
LOWER-BOUND(A, x)
 1  if x > A[A.length]
 2       return error: "not-found"
 3  l = 1
 4  r = A.length
 5  while l < r
 6       m = ⌊(l + r)/2⌋
 7       if A[m] ≥ x
 8            r = m
 9       else l = m + 1
10  return A[r]
```

▷ *Solution 235*

```
CONTAINS-SQUARE(A)
 1  ℓ = A.size
 2  for i = 1 to ℓ − 1
 3       for j = 1 to ℓ − 1
 4            d = 1
 5            while i + d ≤ ℓ and j + d ≤ ℓ
 6                 if IS-SQUARE(A, i, j, d)
 7                      return TRUE
 8                 d = d + 1
 9  return FALSE
```

```
IS-SQUARE(A, i, j, d)
 1  for k = 1 to d
 2       if A[i + k][j] ≠ A[i][j]
 3            return FALSE
 4       if A[i + k][j + d] ≠ A[i][j]
 5            return FALSE
 6       if A[i][j + k] ≠ A[i][j]
 7            return FALSE
 8       if A[i + d][j + k] ≠ A[i][j]
 9            return FALSE
10  return TRUE
```

The complexity of IS-SQUARE(A, i, j, d) is $\Theta(d)$. The complexity of CONTAINS-SQUARE(A) is therefore $O(n^4)$.

▷ *Solution 236*

```
MIN-HEAP-CHANGE(H, i, x)
 1  if x < H[i]
 2       H[i] = x
 3       while i > 1 and H[⌊i/2⌋] > x
 4            swap H[⌊i/2⌋] ↔ H[i]
 5            i = ⌊i/2⌋
 6  elseif x > H[i]
 7       H[i] = x
 8       j = MIN-OF-THREE(H, i)
 9       while i ≠ j
10            swap H[i] ↔ H[j]
11            j = MIN-OF-THREE(H, i)
```

```
MIN-OF-THREE(H, i)
 1  m = H[i]
 2  j = i
 3  if 2i ≤ H.heap-size and H[2i] < m
 4       j = 2i
 5       m = H[2i]
 6  if 2i + 1 ≤ H.heap-size and H[2i + 1] < m
 7       j = 2i + 1
 8  return j
```

The complexity is $\Theta(\log n)$, since in the worst case we would start from a leaf and go all the way up to the root, or we would start from the root and go all the way down to a leaf.

▷ *Solution 237*

```
BST-SUBSET(T_1, T_2)
 1  a = MIN(T_1)
 2  b = MIN(T_2)
 3  while b ≠ NIL and a ≠ NIL
 4       if a.key < b.key
 5            return FALSE
 6       elseif a.key > b.key
 7            b = NEXT(b)
 8       else a = NEXT(a)
 9  if a == NIL
10       return TRUE
11  else return FALSE
```

```
MIN(t)
1  if t == NIL
2       return NIL
3  while t.left ≠ NIL
4       t = t.left
5  return t
```

```
NEXT(t)
1  if t == NIL
2       return NIL
3  if t.right ≠ NIL
4       return MIN(t.right)
5  while t.parent ≠ NIL
   and t == t.parent.right
6       t = t.parent
7  return t.parent
```

The complexity is $\Theta(n)$, since we use MIN and NEXT to effectively iterate over each tree as in a tree walk.

▷ *Solution 238*

This is a classic NP problem. We prove that by showing a polynomial-time verification algorithm. In particular, we show an algorithm VERIFY-CYCLE(G, k, C) that takes an instance of the problem, that is, a graph G and a cycle length k, and a witness cycle C, and verifies that C is indeed a cycle in G of length k.

```
VERIFY-CYCLE(G, k, C)
 1  if C.length ≠ k
 2       return FALSE
 3  for i = 1 to C.length − 1
 4       for j = i + 1 to C.length
 5            if C[i] == C[j]
 6                 return FALSE
 7       if not FIND-NEIGHBOR(G, C[i], C[i + 1])
 8            return FALSE
 9  if not FIND-NEIGHBOR(G, C[C.length], C[1])
10       return FALSE
11  return TRUE
```

```
FIND-NEIGHBOR(G, u, v)
1  Adj = adjacency list of G
2  for w ∈ Adj[u]
3       if w == v
4            return TRUE
5  return FALSE
```

The complexity of VERIFY-CYCLE is $O(n^2)$.

▷ *Solution 239*

This problem is in P. We prove that by showing an algorithm FIND-FOUR-CYCLE(G) that solves the problem in polynomial-time. In particular, the complexity of FIND-FOUR-CYCLE(G) is $O(n^4)$.

```
FIND-FOUR-CYCLE(G)
 1  for a ∈ V(G)
 2       for b ∈ V(G)
 3            if b ≠ a and FIND-NEIGHBOR(G, a, b)
 4                 for c ∈ V(G)
 5                      if c ≠ b and c ≠ a and FIND-NEIGHBOR(G, b, c)
 6                           for d ∈ V(G)
 7                                if d ≠ c and d ≠ b and d ≠ a
                                  and FIND-NEIGHBOR(G, c, d) and FIND-NEIGHBOR(G, d, a)
 8                                     return TRUE
 9  return FALSE
```

```
FIND-NEIGHBOR(G, u, v)
1  Adj = adjacency list of G
2  for w ∈ Adj[u]
3      if w == v
4          return TRUE
5  return FALSE
```

▷ *Solution 240*

Let DP(n) be the number of ways one can express n as a sum of ones, twos, and threes. Then we can immediately write a dynamic-programming recurrence as follows:

$$\text{DP}(n) = \text{DP}(n-1) + \text{DP}(n-2) + \text{DP}(n-3)$$

This is because n can be obtained by adding 1 to all the DP($n-1$) ways one can obtain $n-1$, or by adding 2 to all the DP($n-2$) ways one can obtain $n-2$, or by adding 3 to all the DP($n-3$) ways one can obtain $n-3$. We could then write SUMS-ONE-TWO-THREE(n) recursively as follows:

```
SUMS-ONE-TWO-THREE(n)
1  if n ≤ 0
2      return 0
3  elseif n == 1
4      return 1
5  elseif n == 2
6      return 2 // 1 + 1, 2
7  elseif n == 3
8      return 4 // 1 + 1 + 1, 2 + 1, 1 + 2, 3
9  else return SUMS-ONE-TWO-THREE(n − 1)
       +SUMS-ONE-TWO-THREE(n − 2)
       +SUMS-ONE-TWO-THREE(n − 3)
```

However, the complexity of this solution is most definitely not $O(n)$, since it looks a lot like the recursive version of FIBONACCI, and in fact we can use the same idea to make it efficient. The idea is to compute DP(i) from left to right, starting from the base cases:

```
SUMS-ONE-TWO-THREE(n)
 1  if n ≤ 0
 2      return 0
 3  elseif n == 1
 4      return 1
 5  elseif n == 2
 6      return 2 // 1 + 1, 2
 7  elseif n == 3
 8      return 4 // 1 + 1 + 1, 2 + 1, 1 + 2, 3
 9  else a = 1
10       b = 2
11       c = 4
12       r = a + b + c
13       for i = 5 to n
14           a = b
15           b = c
16           c = r
17           r = a + b + c
18       return r
```

▷ *Solution 241*

The most straightforward solution is one that simply tries all the partitions of $n = a + b$ into two integers a and b greater than 1 Since $n = a + b = b + a$, we can limit the search to $a \leq b$.

```
TWO-PRIMES(n)
1  a = 2
2  while a ≤ n − a
3       if IS-PRIME(a) and IS-PRIME(n − a)
4            return TRUE
5       a = a + 1
6  return FALSE
```

```
IS-PRIME(n)
1  i = 2
2  while i^2 ≤ n
3       if n is divisible by i
4            return TRUE
5       i = i + 1
6  return FALSE
```

The main loop of TWO-PRIME runs for at most $n/2$ iterations, each costing $O(\sqrt{n})$, since the complexity of IS-PRIME(n) is $O(\sqrt{n})$. So, the overall complexity of TWO-PRIME is $O(n\sqrt{n})$.

▷ *Solution 242.1*

ALGO-X(A) returns the lowest category corresponding to the objects in A with a maximal total weight. This is the category c such that there is no other category $k < c$ such that the sum of all the objects in A with category k is higher than the sum of all the objects in A with category c. The complexity of ALGO-X(A) is $\Theta(n^2)$, since the algorithm consists of two nested loops over exactly n.

▷ *Solution 242.2*

We can create a copy of A that is sorted by category, which would allow us to compute the total weight for each category in a single linear pass.

```
BETTER-ALGO-X(A)
 1  B = copy of A
 2  sort B by category
 3  w = −∞
 4  t = B[1].weight
 5  for i = 2 to B.length
 6       if B[i].category == B[i − 1].category
 7            t = t + B[i].weight
 8       else if t > w
 9                 c = B[i − 1].category
10                 w = t
11                 t = B[i].weight
12  if t > w
13       c = B[B.length].category
14  return c
```

▷ *Solution 243.1*

```
MIN-HEAP-INSERT(H, x)
1  H.heap-size = H.heap-size + 1
2  i = H.heap-size
3  H[i] = x
4  while i > 1 and H[i] < H[⌊i/2⌋]
5       swap H[i] ↔ H[⌊i/2⌋]
6       i = ⌊i/2⌋
```

▷ *Solution 243.2*

```
MIN-HEAP-DEPTH(H)
1  i = 1
2  d = 0
3  while 2i ≤ H.heap-size
4       i = 2i
5       d = d + 1
6  return d
```

▷ *Solution 244.1*
ALGO-Y(A) returns the maximal sum of any pair of distinct elements in the input array. If there are less than 2 elements in the array, then the result is $-\infty$. The complexity of ALGO-Y(A) is $\Theta(n^2)$, since the algorithm iterates over all the $n(n-1)/2$ pairs of elements.

▷ *Solution 244.2*
The maximal sum of any two pairs of elements is simply the sum of the two highest values in A. So, we can simply find those two elements and then return their sum:

```
BETTER-ALGO-Y(A)
 1  if A.length < 2
 2       return −∞
 3  i = 1
 4  for k = 2 to A.length
 5       if A[k] > A[i]
 6            i = k
 7  if i == 1
 8       j = 2
 9  else j = 1
10  for k = 1 to A.length
11       if k ≠ i and A[k] > A[j]
12            j = k
13  return A[i] + A[k]
```

▷ *Solution 245*
The problem is in P. As a proof, we show an algorithm MIN-K-SUM(A, m, k) that solves the problem in polynomial time. In fact, this is a greedy problem. In particular, we can answer the question by adding up the highest k values in A. If their total sum is greater or equal than m, then the result is clearly true. Otherwise, the result is clearly false, since there can not be another element that yields a larger sum.

```
MIN-K-SUM(A, m, k)
 1  if k > A.length
 2       return FALSE
 3  B = copy of A
 4  sort B in descending order
 5  s = 0
 6  for i = 1 to k
 7       s = s + B[i]
 8  if s ≥ m
 9       return TRUE
10  else return FALSE
```

The complexity of MIN-K-SUM is $O(n \log n)$.